Communications
Research
Centre

NEURAL AND OPTICAL NEURAL NETWORKS

by

Dr. Kenneth L. Sala
*(Directorate of Optical Communications
and Electrophotonics Technology)*

CRC REPORT NO. 1434

February 1991
Ottawa

This document was prepared for and is the property of the Department of National Defence, Research and Development Branch under Project No. 1410-041LP

Government of Canada
Department of Communications

Gouvernement du Canada
Ministère des Communications

Canada

COMMUNICATIONS RESEARCH CENTRE

DEPARTMENT OF COMMUNICATIONS
CANADA

NEURAL AND OPTICAL NEURAL NETWORKS

by

Dr. Kenneth L. Sala
(Directorate of Optical Communications and Electrophotonics Technology)

CRC REPORT NO. 1434

February 1991
Ottawa

This document was prepared for and is the property of the Department of National Defence, Research and Development Branch under Project No. 041LG14.

91-03503/

ABSTRACT

This report presents an overview of neural and optical neural network technologies which includes detailed descriptions of some of the more important paradigms such as single- and multi-layer perceptrons, backpropagation, Hopfield networks, associative memories, and stochastic learning. Optical neural networks are discussed from the point of view of two classes of implementations, namely, (a), those based upon the use of linear or nonlinear, two-dimensional optical elements such as spatial light modulators or bistable etalon arrays and, (b), those based upon the use of holographic elements. General conclusions regarding the present and future directions for neural and optical neural network research are presented along with specific recommendations for the conduct of an intramural/extramural research program into optical neural networks under the guidance of the Directorate of Optical Communications and Electrophotonics Technology of the Communications Research Center.

RESUME

Ce rapport présente un survol des technologies des réseaux de neurones (optiques et autres), incluant des descriptions détaillées de quelques uns des paradigmes les plus importants, tels que perceptrons à couches simple ou multiple, rétropropagation, réseaux de Hopfield, mémoires associatives, et apprentissage stochastique. Les réseaux de neurones optiques sont discutés dans le contexte de deux classes de mises en oeuvre, soit (a) celles basées sur l'utilisation d'éléments optiques bidimensionnels linéaires ou nonlinéaires tels que les modulateurs spatiaux ou les réseaux d'étalons bistables; ou (b) ceux basés sur des éléments optiques holographiques. Des conclusions générales sur l'état actuel et les directions futures de la recherche en réseaux de neurones et de neurones optiques sont presentées, ainsi que des recommandations spécifiques pour la poursuite de programmes de recherche intra/extramuraux dans le domaine des réseaux de neurones optiques, sous la gouverne de la Direction des technologies de communications optiques et de l'électrophotonique du Centre de recherches sur les communications.

EXECUTIVE SUMMARY

This report presents an overview of the current state of neural and optical neural network theory and technology. Following discussions of the historical development and of the more fundamental characteristics of neural networks in general, detailed descriptions are given of some of the more important paradigms such as single- and multi-layer perceptrons, backpropagation, Hopfield networks, associative memories, and stochastic learning. These paradigms, furthermore, have been shown to have particular suitability for optical implementations. Following a summarization of the intrinsic advantages in the use of optical signals for neural network technology, optical neural networks are then discussed from the point of view of two classes of implementations, namely, (a), those based upon the use of linear or nonlinear, two-dimensional optical elements such as spatial light modulators or bistable etalon arrays and, (b), those based upon the use of holographic elements such as photorefractive crystals which permit a full, three-dimensional implementation of neural network architectures. The report concludes by offering a series of general conclusions regarding the present and future directions for neural and optical neural network research as well as presenting specific recommendations for the conduct of an intramural/extramural research program into optical neural networks under the guidance of the Directorate of Optical Communications and Electrophotonics Technology of the Communications Research Center.

TABLE OF CONTENTS

1. An Overview of Artificial Neural Systems

1.1 Introduction 1
1.2 Outline of Report 3
1.3 Historical Development 4
1.4 Artificial Neural Systems - A Definition and Synopsis .. 9

2. Neural Network Paradigms

2.1 Introduction 14
2.2 Artificial Neurons 15
2.3 Perceptrons 19
2.4 Multi-layer Perceptrons and Backpropagation 22
2.5 Hopfield Nets and Associative Memories 26
2.6 Stochastic and Unsupervised Training Paradigms 31
2.7 Taxonomy of Artificial Neural Systems 34
2.8 Implementation of Artificial Neural Systems 37

3. Optical Neural Systems

3.1 Introduction 40
3.2 Optical Associative Memories and Hopfield Networks 43
3.3 Optical Neural Networks using Holographic Components 48
3.4 Future Research into Optical Neural Networks 52

4. Conclusions and Recommendations

4.1 Conclusions 55
4.2 Recommendations 56

The Neural Network Literature A-1

References REF-1

1.0 AN OVERVIEW OF ARTIFICIAL NEURAL SYSTEMS

1.1 Introduction

The human brain, with an estimated 10^{11} neurons and 10^{15} interconnections, represents nature's foremost archetype of a neural network. Among its many endemic characteristics, the human brain is capable of real time processing of a variety of sensorial inputs along with simultaneous motor control responses, associative storage and recall, short and long term memories, adaptability, generalization, abstraction, and apparently unlimited learning. Yet the brain performs these and other tasks in a package which has a mass of about 1.5 kg, effectively functions continuously for several decades, and operates at a power level of approximately 20 watts.

Research into the development of both a theoretical science and a pragmatic technology of artificial neural systems represents an endeavor to construct "computing" devices which ultimately will be capable of imitating many or all of the characteristics of biological neural systems such as those mentioned above. Fundamentally, artificial neural network research draws its inspiration and motivation from a knowledge of the structure and function of biological neural systems, attempting to model the basic elements and structures of artificial networks after their biological counterparts. It should be noted at the very outset, however, that the aim of artificial neural sytems is NOT a faithful emulation of any biological brain but rather is the design of structures which operate on the same basic and essential principals which govern biological neural networks. Indeed, as our knowledge of biological neural systems expands, it is becoming increasingly apparent that the detailed characteristics of much of the human brain are more complex and sophisticated than was originally held to be true. For example, there are an estimated 100 distinct types of elementary neurons in the human brain with the precise behaviour of any of these basic biological elements proving to be considerably more complicated than was thought to be the case only a few years ago. In reality, it is already correct to say that the correlation between artificial neural systems and biological neural systems is essentially a metaphorical one with respect to the finer details of their respective behaviors. This inexactitude between the model and the imitation should not, however, be seen as some form of fundamental failure or limitation. It is imperative to always keep in mind the meaning of the term "artificial" in discussing any form of artificial neural network. Thus, while the human brain may present itself as the ultimate (perhaps unattainable) paradigm and aim of artificial neural network research, the artificiality of the emulation permits the development or function of artificial neural systems to exploit and exhibit characteristics NOT found in biological ones. To give but one example, artificial neural systems are expected to ultimately operate at MHz or GHz frequencies (with the possibility of THz for all-optical networks) while biological networks effectively run at rates of several hundred Hertz. Such "artificial" advantages may ultimately compensate for any shortcomings in the strict faithfulness of the modelling by artificial neural systems.

The development and pursuit of any new technology raises questions of the need and rationalization for such an activity. This applies to the subject of artificial neural systems even though it should be noted that this field involves to a greater extent novel applications and architectures for information processing (using existing technology) than it does the development or creation of an entirely new technological basis. To answer the question of desirability for artificial neural system research, it is useful to realize that virtually all 'tasks' which are amenable to implementation by automated machinery may be classified as falling into one of two broad categories, namely tasks which are algorithmic in nature and those which are non-algorithmic. An algorithmic task may be defined as any task which can be executed by a finite, deterministic sequence of logical operations - the sequence need not be unique but the end result, the 'solution', must be the same whatever the path followed. Virtually all mathematical manipulations, 'number crunching', calculation of classical scientific paragons (i.e., mathematical models), and logic based processing fall into the algorithmic task classification. An often used simple example is that of long multiplication of real numbers. In conventional computers, this elementary task is performed by a specific 'subroutine' or algorithm based on a prescribed sequence of logical operations involving bit manipulation and logical branching (in contrast to humans who perform this task clumsily based principally on 'look-up' combining with simple recursive rules). Non-algorithmic tasks, on the other hand, are those for which no 'reasonably finite' sequence of logical operations can be defined which would ensure that a correct and unique solution is achieved in a finite time. Principal among such examples of non-algorithmic tasks include the processing of sensorial information such as vision (image recognition, classification), hearing, and speech in humans (but note also the artificial equivalents such as radar classification, voice recognition and synthesis, and robotic image systems). More esoteric examples would include thought processes, reasoning, and emotional response.

Processing technology, in accordance with this classification described above, is itself classified into two categories. The high speed, serial, digital computing technology of today (and this would include the newer, coarse grained parallel architectures) excels at the execution of complex, algorithmic tasks while failing, despite an ever increasing computational throughput speed and capacity, to adequately carry out even some of the simplest non-algorithmic tasks such as image recognition or associative recall. The other 'technology' of biological neural systems ("wetware"), in contrast, excels at real-time processing of massive amounts of sensorial data input with the corresponding control of conscious response. The development of artificial neural systems, in one important sense, is an attempt to develop a science and technology which, by imitation of the design and function of biological neural systems, will present us with a competence and capability which complements almost perfectly our existing and future proficiency at 'number crunching' tasks.

Finally, some remarks are in order concerning the attempt to develop a capacity for solving such non-algorithmic tasks based upon what has come to called "artificial intelligence". Succinctly, this field seeks to 'program' intelligent behavior using existing computing technology by employing concepts such as symbolic manipulation, expert

systems, and 'fuzzy' logic. A basic tenet of this approach is that characteristics of intelligence can be exhibited independently of the specific technology employed to carry out the implementation. In recent years, however, this view has increasingly given way to the belief that some degree of hardware 'emulation' of wetware is necessary to the accomplishment of this task (this issue remains controversial, however). In other words, many researchers now believe that attempts to 'program' intelligence (the "top-down" approach) using artificial 'rules' for information processing on essentially serial, deterministic technology cannot succeed since such an approach is almost diametrically opposite to the design and function of the very 'technology' we seek to emulate, i.e., biological neural systems. An example frequently referred to concerns the isolated task of recognizing a specific (speaker dependent) word. A state-of-the-art supercomputer with a cycle time of approximately 10 nsec can accomplish this task in about 0.5 sec which is almost exactly the same amount of time a human would require for the same task using, however, a neural "computer" with a cycle time of 10 msec. The human brain, despite operating some 10^6 times slower than its supercomputer counterpart, is able to easily match (and vastly outperform in more general tasks) the supercomputer by virtue of its distributed and parallel processing structure. This comparison says, in effect, that the high-speed, serial computing device is simply an inappropriate approach to the execution of such tasks. The design of general, efficient artificial neural systems will permit the execution of this and much more general and purposeful data processing tasks in a way which is immeasurably more competent, practical, and proficient than the complementary technology of digital computation. The reader interested in more details of this debate between the artificial intelligence and connectionist approaches is referred to the articles by Searle [1990] and Churchland and Churchland [1990].

1.2 Outline of Report

This report has two principal tasks; (1), to provide an overview of the status and future capabilities of neural and optical neural networks and, (2), to offer recommendations and general directions for a research program for the development of these technologies in Canada under the guidance and coordination of the Directorate of Optical Communications Technology of the Communications Research Center.

The remainder of this introductory chapter will present a brief history of the development of the theory and technology of artificial neural systems to date followed by a discussion of the various advantages and general characteristics of artificial neural systems.

Chapter two of this report provides a description of the most fundamental topics and paradigms of artificial neural networks (independently of specific technology) including artificial neurons, single-layer networks (perceptrons), multi-layer networks, the backpropagation paradigm, Hopfield nets, associative memories, and stochastic learning. In addition, a discussion is given of the general taxonomy of neural networks as well as the various approaches to the hardware implementation of neural networks. It is stressed

to the reader that this chapter does NOT attempt to provide an exhaustive catalog of the (30+) known artificial neural system paradigms. Such a task is not only beyond the scope of this report but would almost certainly prove counter-productive in that a detailed accounting of artificial neural system taxonomy can easily confuse the non-expert (for such comprehensive discussions, the interested reader can consult any of DARPA [1988], Miller [1987], Simpson [1990], Aleksander [1989], Athale and Davis [1988], Dayhoff [1989], Khanna [1990] or Wasserman [1989]). Moreover, while there exists a plethora of paradigms, in practice only four or five of these models would account for the preponderance of published results in artificial neural systems.

Chapter three examines optical neural networks more specifically including hybrid networks using light as the interconnect medium. A critical comparison of the advantages of optical and electronic technologies is given. Details of selected, specific optical implementations will be presented to provide concrete examples of the methodologies currently being pursued for optical neural networks, using the taxonomic scheme of dividing optical neural network research into those based upon two-dimensional optical elements and those employing (reconfigurable) holographic elements.

Chapter four summarizes a series of conclusions concerning the status and future directions for neural and optical neural research. Finally, explicit recommendations are given for a research program conducted under the guidance of DOCT which would involve an intramural activity at CRC as well as substantial contracting out of specific research proposals to Canadian Universities and, ultimately, to Canadian industries.

Following the body of the report, a brief discussion of the neural network literature is given along with the normal list of references cited in the text of the report.

1.3 Historical Development

Although the field of neural networks and optical neural networks in particular is still very much in its infancy, the scientific roots of this field may be traced back to the nineteenth century and the early parts of this century while many of the most basic aspects of the ideology behind the concepts of artificial neural systems have been discussed since antiquity. Scientists such as Helmholtz, Mach, and Maxwell were fundamentally generalists with interdisciplinary interests, among them the fields of psychology and neurobiology. The earliest known proposal to implement an artificial neural system (using hydraulics) came in 1913 (Russell [1913]) although the actual hardware was never built (the first hardware implementation came only in 1951 by Minsky and Edmonds (see Bernstein [1981]) using a combination of tubes, clutches, and motors).

The technology of the nineteenth and early twentieth centuries simply did not permit a sufficiently detailed characterization of neurophysiology beyond optical microscopy to enable these early workers to formulate quantitative models of neural behavior. More sophisticated measurement technology, particularly electrical and

electronic probe equipment, paved the way to a much more detailed description of individual neuron behavior and neurobiological physiology during the 1920's and 1930's. This in turn led to the first real attempts to model neural systems mathematically by McCulloch and Pitts beginning in the early 1940's (McCulloch and Pitts [1943], Pitts and McCulloch [1947], and see also McCulloch [1965]). The work done by McCulloch and Pitts is commonly regarded as marking the modern origin of artificial neural network research. In addition to devising practical models of neural processors inspired by studies of biological neurons, McCulloch and Pitts were also the first to study the behavior of simple networks composed of such artificial neural processors. However, their approach was to investigate a fixed architecture with prescribed values for the interconnection weights by studying the variation in the outputs of the network as the input and/or the interconnection weights were varied. No attempts were made to "train" the networks, i.e., to adjust the weights in order to produce a prescribed set of output values.

In 1949, the psychologist Hebb (Hebb [1949]) enunciated a qualitative rule describing the neural mechanism by which learning behavior was governed. This principal, now called 'Hebbian learning', was quantified and enlarged upon by later workers and, in its many variations, underlies the learning algorithms used by nearly all the major artificial neural system paradigms in existence today. The apparently subtle but fundamentally important perception that interconnection weights could be adjusted to train an artificial neural system at some prescribed task (the concept of learning with its attendant characteristics of abstraction and generalization) led to many attempts to build or model such systems throughout the 1950's and 1960's by researchers such as Minksy [1961, 1969], Rosenblatt [1958, 1962], Uttley [1966], Widrow [1959, 1987], Widrow and Hoff [1960], Steinbuch [1961,1965], Grossberg [1968a, 1968b, 1969], and many others. Many of the these early efforts proved exciting and fruitful and artificial neural system research appeared to offer great promise for the development of larger, more general and powerful artificial neural systems.

Rosenblatt in particular played a central role during this time by formally introducing the first specific neural network paradigm, the perceptron, in 1958 (Rosenblatt [1958]) and by providing a rigorous, quantitative theory for this elementary artificial neural system (Rosenblatt [1962]). Throughout the late 1950's and 1960's, many other workers systematically studied perceptrons (the term 'perceptron' by common usage is understood, unless otherwise modified, to denote 'single-layer perceptron'). Relatively modest perceptrons could be 'taught' to perform tasks which could either not be done at all or only with significant difficulty by then-current EDP technology (see Block [1962] and Widrow [1987]).

However, by the mid 1960's, it was explicitly noted that the perceptron suffered certain fundamental limitations in its ability to represent general problems, limitations which essentially destroyed its usefulness as a general purpose system. Although the most telling "condemnation" is rightfully attributed to the publication of Minsky and Papert [1969], it is fitting to note that other workers had also pointed to such limitations. The most serious of these concerned the ability of the perceptron to represent only so-

called'linearly separable' problems. As figure 1.1 indicates, the proportion of linearly separable problems falls dramatically as the number of inputs increases, making the (single-layer) perceptron truly useful only on a small scale. In particular, Minsky and Papert showed that a two element perceptron was inherently unable to perform the logical XOR (exclusive OR) function on the inputs. Furthermore, Minsky and Papert had also predicted that the perceptron's shortfalls would not be overcome by multi-layer architectures with so-called 'hidden layers'. This latter prediction was eventually proven to be utterly incorrect when it was shown that a three- (or more) layer perceptron could represent a 'decision region' in multi-dimensional space of arbitrary shape and convexity (this important capability is explained and illustrated in detail in Lippman [1989], Khanna [1990], Simpson [1990], and Wasserman [1989]).

Such basic criticism of the technique resulted in the near annihilation of the still-fledgling field of neural network research. Although many workers believed heuristically and intuitively that such limitations would be definitively eliminated if properly trained MULTI-LAYER perceptrons could be designed, no method was known at the time for systematically training inner or 'hidden' neural layers in such a way as to reasonably insure convergence of the network to a stable local or global minimum in the error space. In short, no paradigm and general training algorithm existed beyond the simple one-layer perceptron. With the present profusion of network paradigms, it can appear perplexing to newcomers to the field that no such algorithms were known since many of the present algorithms appear so 'obvious'. Such an appearance, however, stems from using 20-20 hindsight. Within a very short period of time, many workers (along with virtually all research funds) abandoned the field, most turning their attention to the new field of 'artificial intelligence'. Neural researchers often refer to the decade of the 1970's as the "dark ages".

However, a handful of 'true believers' (see the appendix in Simpson [1990] for a brief discussion of the chronology and contributions of several of the key researchers during this period) persisted in their work on neural networks and several largely distinct learning paradigms and artificial neural system architectures emerged during the 1970's. By the mid 1980's, a general paradigm for the training of multi-layer neural networks known as 'backpropagation' emerged (it is one of the ironies of this field that this solution to deterministic training of hidden layers, the so-called "credit assignment problem", although actually published by Amari (Amari [1967]) some two years prior to the book by Minksky and Papert, went almost completely unnoticed as did similar solutions offered by Werbos [1974] and Parker [1982]). This technique proved to be enormously useful and practical and remains one of the best and most general algorithms available today. Since that time, even more algorithms have appeared, some as variations or enhancements of the backpropagation method while others represent distinct approaches to this task. These and other key developments such as the Hopfield network led to a rekindling of interest in neural network research although substantive increases in funding for this field would not begin until the mid 1980's. Figure 1.2 gives a tabular summary of the historical development of artificial neural systems. The latter half of the 1980's has witnessed a prodigious growth in the number of papers and books published concerning artificial

Linear Separability versus Number of Inputs

n	2^{2^n}	No. of linearly separable functions	%
1	4	4	100
2	16	14	87.5
3	256	104	40.6
4	65,536	1882	0.029
5	4.3×10^9	94,572	0.0022
6	1.8×10^{18}	5,028,134	2.8×10^{-11}

Figure 1.1 - Fundamental limitation on the applicability of the single-layer perceptron

Historical Milestones in the Development of Artificial Neural Systems

- 1913 - First ever proposal for hardware model of artificial neural system (Russell).
- 1943 - McCulloch and Pitts publish first work on theory of artificial neurons.
- 1949 - Hebb describes learning mechanism principles for neural networks.
- 1958 - Rosenblatt first describes the single-layer perceptron.
- 1950's & 1960's - Many researchers investigate new learning paradigms and applications of elementary neural networks.
- 1964 - Grossberg begins studies of self-organizing networks, leading eventually to the adaptive resonance theories.
- 1969 - Book "Perceptrons" by Minsky and Papert leads to near abandonment of neural network research.
- 1971 - Kohonen begins research into randomly connected paradigms and self-organization algorithms.
- 1970's - The "dark ages" of neural network research; only a handful of researchers actively pursue the subject. Nevertheless, many advances made during this time in mathematical foundations and new conceptual models.
- Early 1980's - First appearance of commercial companies dealing exclusively with neural network hardware/software.
- 1982 - Hopfield proposes novel class of recurrent networks and associative memories.
- Mid 1980's - Rapid resurgence of interest in neural network research. First optical implementations reported.
- 1986 - Explicit description of the backpropagation algorithm.
- Late 1980's - Neural networks becomes a major research field. Many reports of hardware implementations. Rapid proliferation of published material, conferences, and new journals.

Figure 1.2 - Historical development of artificial neural systems

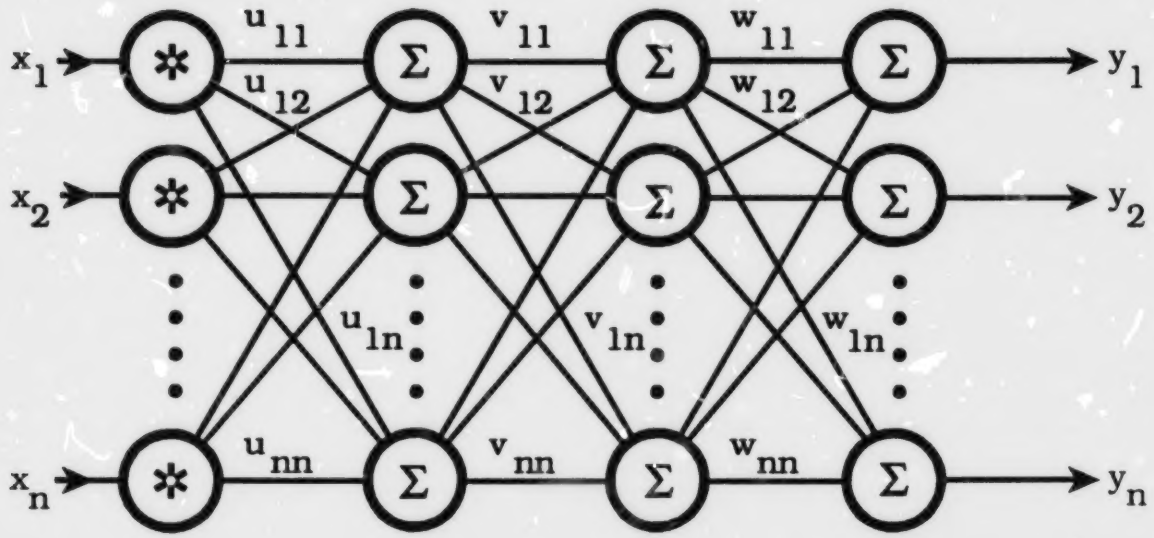
neural systems along with an ever increasing number of conferences and workshops dealing exclusively with one or more aspects of artificial neural system theory and experiment. Several commercial companies in the USA, Canada, and in Europe arose during the 1980's (see Rosenfeld [1987]) whose principal and often sole product is artificial neural system software and/or hardware for research and 'real-world' applications.

A significant aspect of the rekindling of interest and activity in artificial neural system research during the 1980's is that it resulted not only from a growing list of accomplishments from the field but also from a growing sense of disenchantment and disaffection for the field of artificial intelligence. The latter had certainly failed to live up to many of the promises made by its proponents. As significantly, a growing number of workers both inside and outside of the artificial intelligence community became convinced that many of the long-term objectives of that field would not and could not be accomplished solely through the 'software' approach, i.e., through algorithmic design independently of the hardware and its architecture (the neural network or connectionist approach to building a truly 'intelligent' computer is often called the "bottom-up" (or inside-out) approach while strict AI represents the "top-down" (or outside-in) approach). A consensus has grown throughout the 1980's that ultimately true artificial intelligence can only be achieved by hardware and 'software' which is modeled after biological neurosystems. Many others, it should be noted, believe that, in the longer term, it will be the marriage of neural network technology with certain AI concepts which will lead to major achievements in artificial 'brains'.

1.4 Artificial Neural Systems - A Definition and Synopsis

A general but rigorous definition of an artificial neural system which has become widely accepted as the de facto definition by the neural network community is that given by Hecht-Nielsen [1988] as follows (and summarized in figure 1.3):

> A neural network (artificial neural system) is a parallel, distributed information processing structure consisting of processing elements (which can possess a local memory and carry out localized information processing operations) interconnected together with unidirectional signal channels called connections. Each processing element has a single output connection which branches ("fans out") into as many collateral connections as desired (each carrying the same signal - the processing element output signal). The processing element output signal can be of any mathematical type desired. All of the processing that goes on within each processing element must be completely local; i.e., it must depend only upon the current values of the input signals arriving at the processing element via impinging connections and upon values stored in the processing element's local memory.



A "Neural Network" is a parallel, distributed information processing structure consisting of processing elements interconnected together with unidirectional signal channels called connections. All processing is completely local, i.e., it is done using only current values to the processor and its local memory.

Figure 1.3 - An abbreviated definition of an artificial neural system

The two most critical and significant descriptors in this definition are the adjectives 'parallel' and 'distributed'. It is the combination of these characteristics which fundamentally distinguishes a neural network ("fine grain parallelism") from what are referred to as massively parallel computer architectures ("coarse grain parallelism") such as the Connection Machine and the Cray X/MP. Coarse grain parallelism retains much of the underlying structure of serial (von Neumann) computers (such as separate CPU and memory hardware) and essentially represents an attempt better termed as concurrent, multi-processor computing in order to dramatically improve computational throughput (in truth, the architectural changes are more involved than this alone would indicate - see the discussions by Fox and Messina [1987], Peled [1987], and Soucek [1989]). It is worthy of note that exists whole classes of problems which are intrinsically unamenable to this approach to parallel co-processing although much is yet to be learned in the task of parallel programming and optimal use of such novel computer architectures.

In a true neural network, in contrast, the distributed AND parallel nature of the processing elements and their interconnections means that the input data itself is processed on a 'bit-by-bit' basis in parallel and that each processor runs independently of each other processor, relying entirely on its input values and the values of the 'synaptic' weights on its immediate interconnects. These weights serve as both the memory of the network as well as the encoding of the network 'algorithm'. There is no single central processor or separation drawn between memory and processing distribution elements. Figure 1.4 offers an abbreviated list of advantages endemic in the applications of neural networks.

The first of those points listed in figure 1.4, learning, is perhaps one of the most fundamental and advantageous characteristics of artificial neural systems. Not only does learning offer the possibility of using artificial neural systems where no deterministic model or algorithm is known a priori, but it is often true that an artificial neural system can learn some specific task (e.g., speech recognition) with only a fraction of the time, hardware, effort, and cost involved in accomplishing the same task via programming on conventional computers.

The ability of artificial neural systems to generalize is also one of the most attractive features of such systems since this is a characteristic closely associated with biological neural systems and quite uncharacteristic of conventional algorithmic programming. Many artificial neural systems have demonstrated the ability to provide meaningful ("correct") solutions when presented with input data which was well outside of the domain of the training data set. This is generally believed to be the result of the fact that artificial neural systems, when correctly configured and trained, form internal representations of the training data set and the network task which allow the network to subsequently respond in a meaningful and sensible way to new data.

The three characteristics of fault tolerance, abstraction, and realtime response are technical advantages to artificial neural systems which can be vital in many real applications and operating environments. Fault tolerance results from the distributed

Advantages of Artificial Neural Systems

- LEARNING
  ANS's can be trained to perform tasks for which no a priori model or algorithm is known. Learning can be used to modify a network's behaviour as inputs are changed, i.e., networks can be made adaptable.

- GENERALIZATION
  ANS's are able to "generalize" from internal models to provide acceptable solutions for inputs outside those used for network training.

- FAULT TOLERANCE
  Since ANS's are parallel distributed systems, they are inherently tolerant to the loss of or small variations in small numbers of processors or interconnects.

- ABSTRACTION
  Many ANS's exhibit the ability to provide exact solutions when the input set is noisy or distorted.

- REALTIME RESPONSE
  Since ANS's are fine-grained, parallel systems, they are capable of processing sensorial inputs in realtime. Larger neural networks, particularly optical neural networks, will be capable of unprecedented processing throughputs.

Figure 1.4 - Advantages of Artificial Neural Systems

nature of artificial neural systems and enables an artificial neural system to sustain some damage or loss of processors or interconnects without catastrophic failure of the artificial neural system as a whole. Abstraction refers to the ability of most artificial neural systems (particularly those operated in a recurrent or feedback mode) to recall an exact output vector or 'solution' when presented with a distorted and/or noisy input vector. Finally, like its wetware inspiration, artificial neural systems are intrinsically capable of high data throughout by virtue of processing the input data along many independent channels. With the growth in the state-of-the-art size of actual artificial neural systems, the ability to process in realtime ever more complex sensorial inputs will grow to the point, hopefully, where realtime processing comparable to that of the human brain may be realizable (perhaps modularly). More comprehensive discussions of the characteristics of artificial neural systems may be found in Aleksander [1989], DARPA [1988], Dayhoff [1989], Grossberg [1988], Simpson [1990], Kohonen [1988], Miller [1989], and Wasserman [1989]. A fascinating, speculative discussion of the long term possibilities of the development of technological "intelligence" is given in Moravec [1988].

2.0 NEURAL NETWORK PARADIGMS

2.1 Introduction

A general artificial neural network requires the specification of three components:

(1) a model for the processing element or artificial neuron (keeping in mind that some networks employ more than one type of artificial neuron);
(2) the architecture or topology by which the artificial neurons are interconnected including any feedback (without loss of generality, neural networks are viewed as connected layers of processing elements); and
(3) an algorithm which quantitatively specifies how the interconnection weights will be adjusted as training data is presented to the network along with a criterion to decide when the network is fully trained.

Implicit in this process is the assumption of the existence of a suitable set of prescribed input/output data which is used to train the network (i.e., the learning stage) but which is irrelevant to the operation of the network once the training is completed.

A large number of diverse and varied approaches to each of the components above, particularly the architecture and learning algorithms, has led to the creation of a large number of distinct paradigms for artificial neural networks (although several of these paradigms are, in principal, subsumed by other, more general designs). It is neither practical or particularly useful to attempt to describe each of these paradigms since, in practice, many of them have received little attention or generated noteworthy applications. Instead, the single- and multi-layer perceptron networks will be described as representative models along with the backpropagation algorithm for the training of multi-layer networks. This paradigm remains one of the most widely used and practical training algorithms. The basic architectures and principles of operation of Hopfield and associative memory networks will be described along with a discussion of statistical training methods. Following this, an overview of artificial neural network taxonomy will be presented along with a brief discussion of some of the reasons behind the proliferation of artificial neural network paradigms. Finally, a discussion will be given concerning the various methods employed to date for the implementation of artificial neural systems. Much more comprehensive treatments of virtually all the present network paradigms and learning algorithms are given by Simpson [1990], Khanna [1990], Miller [1987], and DARPA [1988].

2.2 Artificial Neurons

At the very basis of all artificial neural systems is the model for the elementary processing element of the network, the artificial neuron. Generally speaking, the function of the artificial neuron is threefold:

(1) to produce a weighted sum of all inputs it receives;
(2) to transform this sum through some (normally nonlinear) operation such as a threshold function or sigmoidal activation function; and
(3) to send this output value to all those neurons in a successive layer to which it is connected.

Typical biological neurons appear to operate mostly on an "all-or-nothing" basis, i.e., the neuron 'fires' if and only if the summed inputs to it exceed some threshold value. This model was the basis of the earliest work by McCulloch and Pitts and is still is use today although most workers use an analog artificial neuron for which the transfer or activation function is some analytic mathematical function.

Figure 2.1 illustrates the basic geometry and terminology for a single artificial neuron. In mathematical terms, the various input signals x_i are summed, each multiplied by the corresponding weight value w_{ij} (the 'synaptic weight') to produce the input value 'net',

$$net_j = \sum_{i=1}^{n} x_i w_{ij} \tag{2.1}$$

This value provides the input argument for the activation or transfer function 'F' of the artificial neuron producing a value 'out' given by

$$out_j = F(net_j + T_j) = F(\sum_{i=1}^{n} x_i w_{ij} + T_j) \tag{2.2}$$

where T_j is a (independent) threshold value for the jth neuron. The most commonly employed transfer functions are:

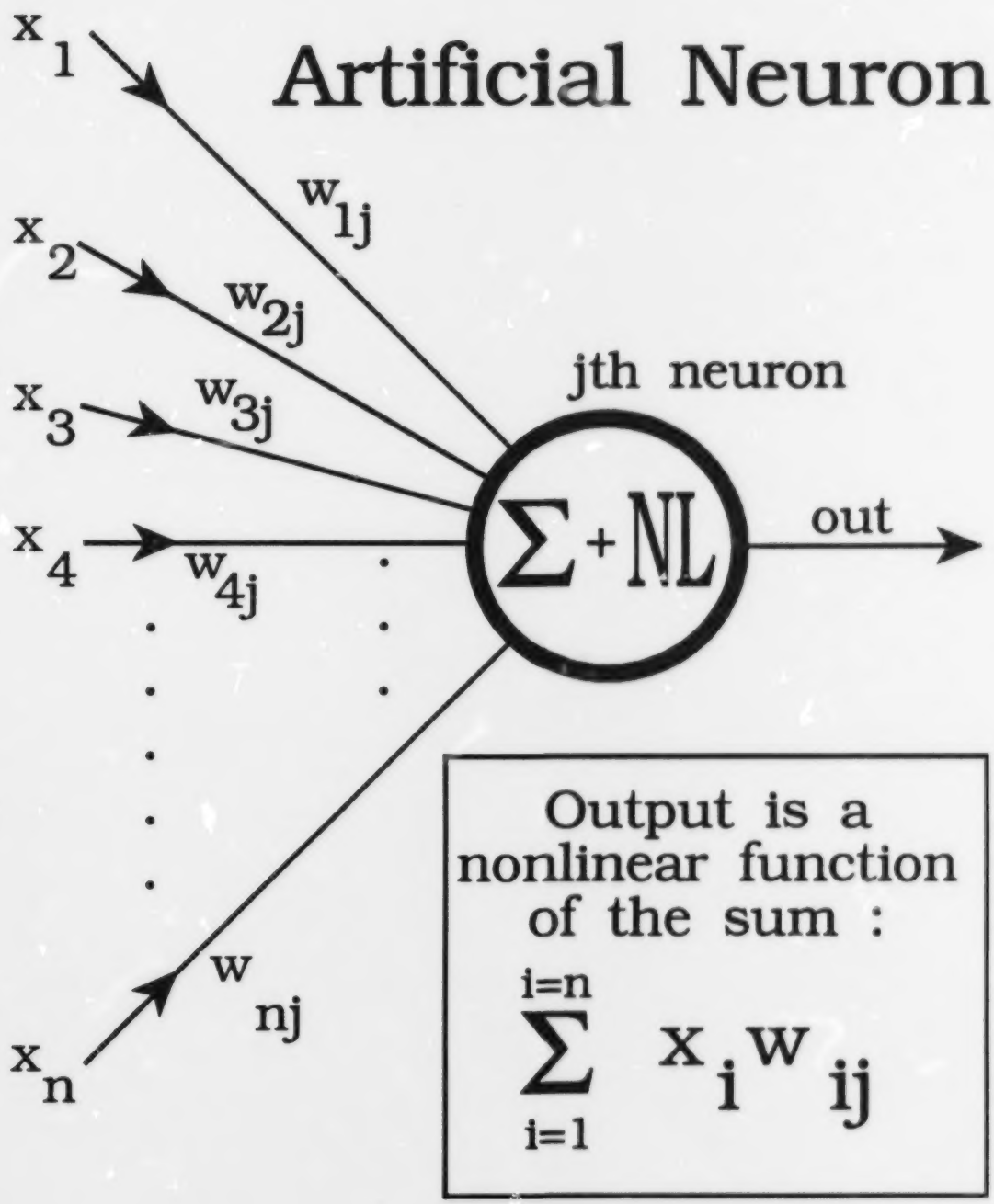


Figure 2.1 - Basic model of an artificial neuron

$$
\begin{aligned}
&(a)\ F(x) = x \quad ; \quad |x| \le \beta \\
&\qquad\ = +\beta \quad ; \quad x > \beta \\
&\qquad\ = -\beta \quad ; \quad x < -\beta \\
&(b)\ F(x) = sgn(x) \\
&(c)\ F(x) = (1+e^{-x})^{-1} \\
&(d)\ F(x) = \tanh(x) \\
&(e)\ F(x) = x^2/(1+x^2) \ ; \ x \ge 0 \\
&\qquad\ = 0 \quad ; \quad x < 0
\end{aligned}
\tag{2.3}
$$

and are illustrated in figure 2.2. The threshold or 'hard limiter' model corresponds most closely to "all-or-nothing" neural behavior and is often the model used for nets with discrete neural outputs (noting that it is still necessary for the neuron to perform the weighted sum of its inputs in an analog fashion).

The use of the various types of sigmoidal functions is required in order to ensure that output values do not become exceedingly large when summed input parameters themselves become large. In the absence of such "squashing" functions, both inputs and outputs in a network could easily become extremely large for a subgroup of interconnections with the result that various interconnection weights would tend to become extremely large in absolute value. These interconnections would then totally dominate and determine the behavior of the overall network. Viewed in another manner, the sigmoidal transfer function has a derivative (related to a gain factor in the changing of interconnect weights) which is largest about zero (where weights will be small) and smallest at the extremities (where weights will tend to be large in absolute value), i.e., the sigmoidal shape serves as a 'gain limiter' in the iterative adjustments of the interconnect weights and so serves to prevent 'runaway' or divergent values of the weights during the training process.

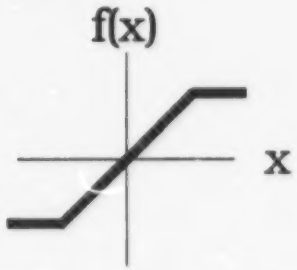
Other special types of artificial neurons have been reported, frequently with transfer characteristics defined on a layer-by-layer basis rather than for a single neuron. The Kohonen neuron, for example, acts as a hard limiter but is constrained in such a way that, for any given input, one and only one neuron in the Kohonen layer will fire. The Grossberg neuron is one for which the transfer function is strictly linear (non-truncated domain) and is frequently used as the output layer of a multi-layer network where it serves primarily as a distribution layer.

The various elements of the artificial neuron are modelled (loosely) after the components of the biological neuron, e.g., the interconnects correspond to the axon-dendrites of the biological neuron while the weights correspond to the synaptic barrier. For a more detailed description of the biological neural system, the reader is referred

Neural Transfer Functions

Truncated linear ramp

$f(x)=x$ for $|x| < \beta$



Step threshold function

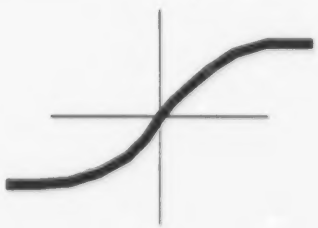
$f(x) = \mathrm{sgn}(x)$

Logistic function

$f(x) = (1 + e^{-x})^{-1}$

Hyperbolic tangent

$f(x) = \tanh(x)$



Augmented ratio of squares

$f(x) = x^2/(1 + x^2)$ for $x>0$

$= 0$ otherwise

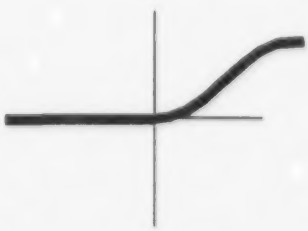


Figure 2.2 - Common transfer functions for artificial neuron models

to Wasserman [1989 - Appendix A] or the DARPA review (DARPA [1988] - Chapter 15). In addition, it should be noted that the description given here for the artificial neuron is a elementary one. Much more complete (but complex) theories of neuronal behavior may be found in Hoppensteadt [1986] and Koch and Segev [1989].

2.3 Perceptrons

The perceptron was the invention of Rosenblatt (Rosenblatt [1958]) and has been extensively studied by several groups since the late 1950's. The term "perceptron" is generally used to indicate a single-layer, feed-forward neural net (multi-layer perceptrons will be discussed in the following section). The neuron model may be analog with a nonlinear transfer function, or discrete with an "all-or-nothing" characteristic. Figure 2.3 illustrates such a perceptron architecture and it will be assumed in the following that the input/output values are continuously variable (the general case).

Each component y_j of the output vector is calculated exactly as prescribed in the previous section, i.e.,

$$y_j = out_j = F(net_j) = F(\sum_{i=1}^{n} x_i w_{ij}) \tag{2.4}$$

so that the relationship between the input and output vectors may be expressed compactly in the form

$$Y^t = F(XW) \quad , \tag{2.5}$$

where $\mathbf{Y}^t$ is the column vector (the "t" denoting a transpose)

$$Y^t = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix} \quad , \tag{2.6}$$

X is the row vector

The Single-Layer Perceptron

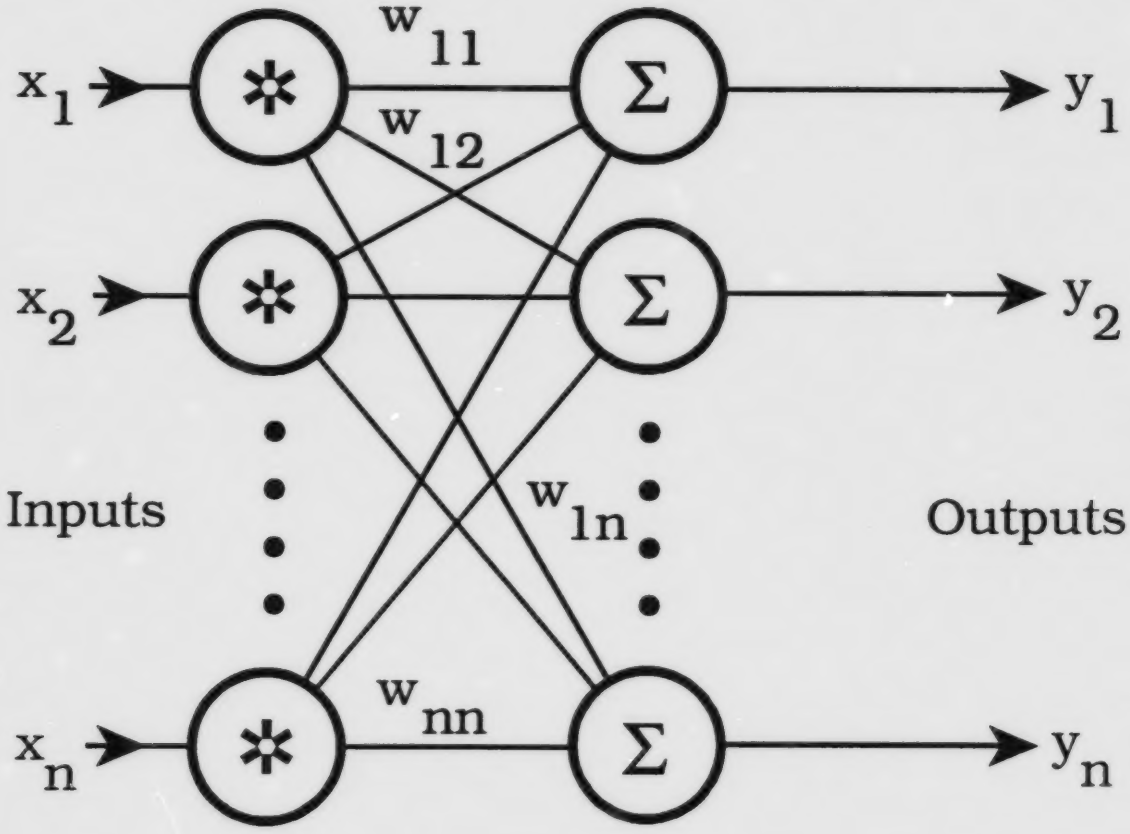


Figure 2.3 - The architecture of the single-layer perceptron

$$X = (x_1 x_2 \cdots x_n) \quad , \tag{2.7}$$

and W is the interconnection weight matrix

$$W = \begin{pmatrix} w_{11} & w_{12} & \cdot & \cdot & w_{1n} \\ w_{21} & w_{22} & \cdot & \cdot & w_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ w_{n1} & w_{n2} & \cdot & \cdot & w_{nn} \end{pmatrix} \tag{2.8}$$

The training of the neural network is an iterative procedure involving pairs of matched input/output vectors $\{A_r, B_r\}$, r = 1,2, ,M. Qualitatively, the algorithm proceeds as follows:

(1) Apply one of the input vectors A_r and calculate the output vector Y_r;
(2) If $Y_r = B_r$, then proceed to step (5);
(3) Calculate the difference for each component of $(B_r - Y_r)$ and adjust the corresponding weights to each output neuron;
(4) Return to step (1);
(5) Repeat steps (1) to (4) for successive members of the training set $\{A_r, B_r\}$ until all the calculated outputs match (to some prescribed degree of accuracy) the corresponding given output vectors.

To put this in quantitative terms, let the difference between the expected and calculated outputs for neuron j and training vector pair r be denoted by

$$\delta_j = (B_j - Y_j) \tag{2.9}$$

The weights are then adjusted according to the rule

$$w_{ij}(t+1) = w_{ij}(t) + \alpha A_i \delta_j \tag{2.10}$$

where α denotes a learning rate coefficient ($0 < \alpha << 1$) and "t" denotes a training counter, i.e., $t \rightarrow t+1$ after each training session is completed for one training vector pair (note that the subscript "r" has been suppressed in the writing of equations (2.9) and (2.10) - it is implicitly understood that such weight adjustment equations are applied successively and iteratively to each member of the training set $\{A_r,B_r\}$). Equation (2.10) expresses quantitatively (one form of) the principle of Hebbian or correlation learning in that the synaptic strength (weight) is increased when BOTH the source (the input) and the neuron are activated. Equation (2.10) is often referred to as the 'delta rule' and modifies the interconnection weights appropriately for outputs of either polarity and for both continuous and binary inputs and outputs.

The procedure just described does not address the question of "matching" the size of the training set to the neural network (or vice versa). There is no rigorous or general answer to this problem although empirical studies have led to the rule of thumb

$$M > \frac{n^2}{\epsilon} \tag{2.11}$$

where ϵ is the allowed fraction of errors on the training set (typically, $\epsilon < 1/8$). Too small a set of training vectors for a given net will "undertrain" the network with the result that more errors will be present and the network will not generalize reliably. On the other hand, too large a training set for a given net will "overtrain" the network and may again increase the errors and make generalization by the network questionable. In addition, too large a training set will nearly always lead to excessively long training times and poor convergence. These points are discussed in more detail by Baum and Haussler [1988] and by Hinton [1989].

2.4 Multi-layer Neural Networks and Backpropagation

A three-layer, feed-forward neural network (perceptron) is illustrated in figure 2.4. The middle layer is called 'hidden' since neither its outputs or inputs are directly connected to either the network input or output nodes. The difficulty (and one which stymied many researchers for several years) was to devise a sound (i.e., stable and convergent) algorithm by which the hidden layer weights could be adjusted. Rumelhart, Hinton, and Williams [1986a,b] are commonly credited with explicitly describing the technique known as backpropagation although, once their work was published, it was realized that the essentials of this method had been previously described by Amari [1967], Werbos [1974], and Parker [1982].

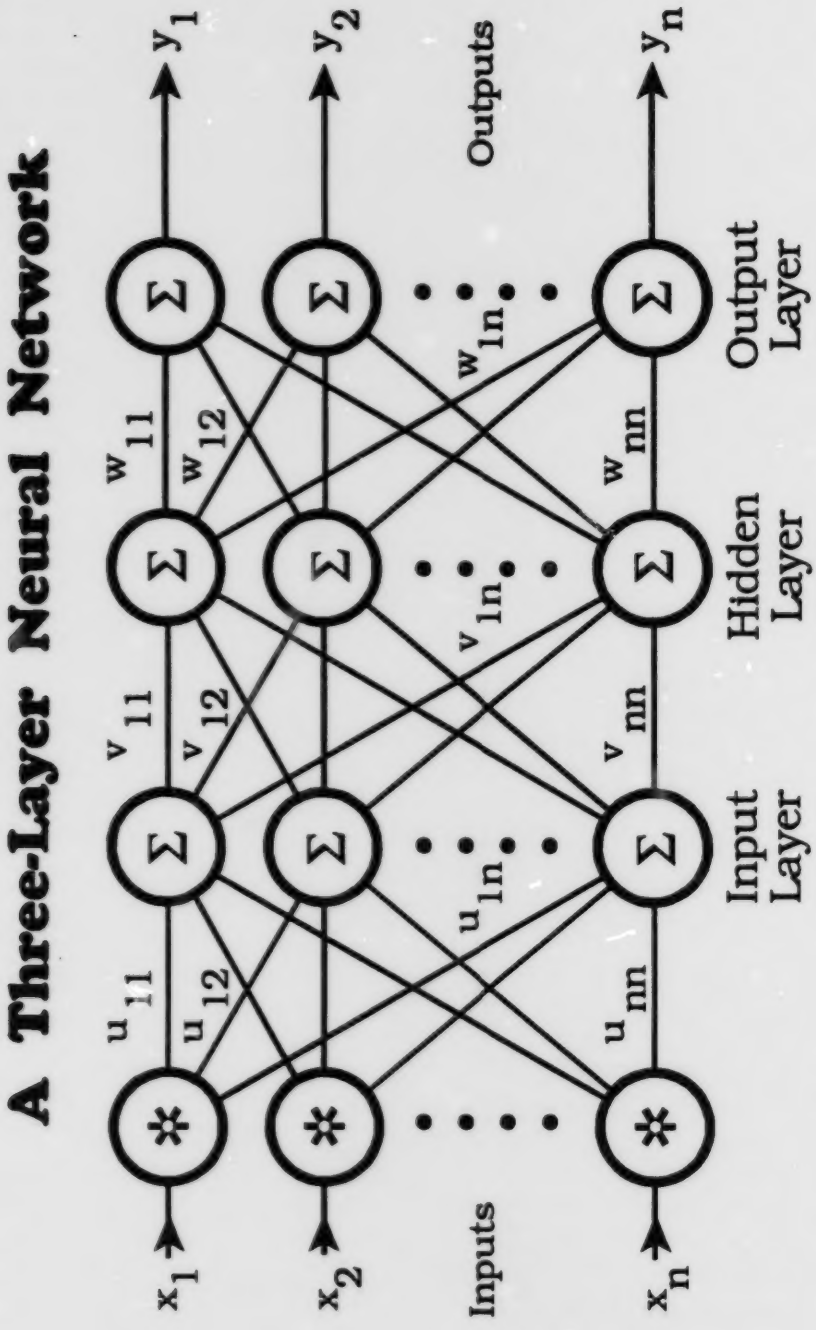


Figure 2.4 - A three-layer perceptron

The backpropagation algorithm consists of the following iterative steps:

(1) As for the case of the single-layer perceptron, begin by applying one of the input vectors $\mathbf{A}_r$ to the network and calculate the output vector $\mathbf{Y}_r$ on the forward pass through the network (i.e., from the input nodes to the output nodes);
(2) By comparison of the expected output vector $\mathbf{B}_r$ corresponding to $\mathbf{A}_r$ with the calculated value $\mathbf{Y}_r$, an error value $\delta_{j,w}$ may be calculated for each of the output neurons (the subscript "w" is used here simply to reference the w or output neuron layer);
(3) Adjust the weight coefficients w_{ij} exactly as for the single-layer case;
(4) Using these adjusted weights and the values of $\delta_{j,w}$, calculate new error coefficients $\delta_{i,v}$ to be used in the adjustment of the hidden layer weights v_{hi};
(5) Repeat steps (3) and (4) for each hidden layer successively back to and including the input layer of neurons; and
(6) Repeat steps (1) to (5) for successive members of the training set $\{\mathbf{A}_r,\mathbf{B}_r\}$ until no further adjustments to any of the weights are necessary.

The key step in this algorithm is clearly step (4) since it provides the necessary error or delta values needed to adjust the hidden layer weights according to an equation such as equation (2.10).

Denoting the outputs from each of the layers as $y_{i,c}$ where "c" = u, v, or w and i = 1,2, ,n, the training of the output layer weights is given by the equation

$$w_{ij}(t+1) = w_{ij}(t) + \alpha \delta_{j,w} y_{i,v} \tag{2.12}$$

where

$$\delta_{j,w} = y_{j,w}(1 - y_{j,w})(B_j - y_{j,w}) \tag{2.13}$$

and α is the learning rate coefficient. The particular form for $\delta_{j,w}$ given above corresponds to the assumption of a logistic function (see figure 2.2) as the nonlinear transfer function for the artificial neurons, i.e.,

$$y_{j,w} = \frac{1}{1 + e^{-net_{j,w}}} \tag{2.14}$$

which has the simple derivative

$$\frac{d y_{j,w}}{d net_{j,w}} = y_{j,w}(1 - y_{j,w}) \qquad (2.15)$$

and it is this derivative which enters into the expression for $\delta_{j,w}$ given in equation (2.13). The logistic function is the transfer function most commonly used in neural networks which learn by backpropagation and is adopted here for that reason. No loss of generality ensues since this algorithm works equally well with any sigmoidal activation function.

To carry the training further, i.e., to the first hidden layer, we calculate new error values for each of the neurons in the hidden layer 'v' according to

$$\delta_{i,v} = y_{i,v}(1 - y_{i,v}) \sum_{j=1}^{n} \delta_{j,w} w_{ij}(t+1) \qquad (2.16)$$

and

$$v_{li}(t+1) = v_{li}(t) + \alpha \delta_{i,v} y_{l,u} \qquad (2.17)$$

Exactly the same process is repeated for the first (input) layer whence the entire procedure is repeated beginning with the forward pass and a new training vector pair. The conceptual crux of this overall procedure corresponds to using the original error values $\delta_{j,w}$ as "input" values and then to operate the network in reverse by "backpropagating" these values from the output layer to the input, adjusting each layer of weights in succession. Note in equation (2.16) that it is the adjusted coefficients $w_{ij}(t+1)$ which are used to calculate the delta values for the hidden layers as the calculation proceeds from the output layer to the input layer, i.e., the reverse pass.

The introduction of this learning paradigm in 1986 led directly to an intensification of research efforts by many individuals and to several promising applications of multi-layer networks in speech processing and pattern recognition and classification problems. It is important to realize that the three-layer perceptron is a true general neural network in terms of its representation ability in contrast to the case for the single- (or two-) layer network. The existence of the hidden layer allows the three-layer network to form 'decision regions' (in hyperspace) of arbitrary shape and convexity (recall that the single-layer perceptron could represent only linearly separable problems). A more detailed description of this representation capability of perceptrons may be found in Lippmann [1987], Khanna [1990], Simpson [1990], and Wasserman [1989].

Although the backpropagation algorithm is a fairly robust one which works reasonably efficiently for many type and configurations of feed forward, multi-layer networks, several important variations to the algorithm have been put forward which enhance its operation in many instances (including its application to recurrent networks, see, e.g., Pineda [1988]). Rumelhart, Hinton, and Williams [1986a] described a method, called "momentum", in which the weight adjustment equations (2.12) and (2.17) include an additional term which is proportional to the amount of the previous weight change for that coefficient. The effect appears to be a 'smoothing' of the variations in the networks weight coefficients as a stable solution is reached. Sejnowski and Rosenberg [1987] have described a similar method based upon exponential smoothing of the weight adjustments which has proved superior in certain applications. Parker [1987] modified the backpropagation algorithm to include second order derivatives to produce a more accurate estimate of the required weight adjustments. He further showed that this improvement is optimal in that including higher-than-second-order derivatives would not enhance the adjustment estimates. Stornetta and Huberman [1987] incorporated a simple but efficacious modification to backpropagation by changing the dynamic range of the activation function from [0,1] to [-½,+½] by the addition of a bias term. The times required to train several networks by backpropagation were reduced from 25% to 50% by this elementary modification.

Despite the success of the backpropagation paradigm, it not a panacea for all cases. Difficulties with network paralysis during learning, temporal instabilities, and failure to reach a 'global minima' have been reported. Further discussion and details of the strengths and weaknesses of this algorithm are given in Simpson [1990], Khanna [1990], Stornetta and Huberman [1987], Pineda [1988], Rumelhart, Hinton, and Williams [1986a], Parker [1987], Wasserman [1989], and Hinton [1989].

2.5 Hopfield Nets and Associative Memories

The single- and multi-layer perceptron networks, trained with the aid of the backpropagation algorithm, are excellent examples of feedforward networks which are deterministically trained. A tacit characteristic of such networks is their stability, i.e., any such trained network will produce one and one output for a given input vector (whether that output is the "correct" one notwithstanding). Recurrent networks, in contrast, explicitly have some degree of feedback from the network outputs (in some cases, outputs from intermediate stages) to the input layer. Such networks operate by iteratively recalculating the output until its value "settles" to one stable vector. Although recurrent networks can give rise to problems of stability, the range of their behavior and capabilities is considerably greater than non-recurrent networks.

One of the best known and most widely studied paradigms for recurrent networks is actually a class of networks which have come to be known as Hopfield networks. Hopfield ([1982, 1984, 1987] and Tank and Hopfield [1987]) published several papers on these networks in the early- to mid-1980's and their work rekindled widespread interest

in artificial neural systems in general. Figure 2.5 illustrates the basic architecture of a single-layer, recurrent Hopfield network (this depiction is somewhat different from that usually drawn for a Hopfield network but is functionally equivalent and allows a ready comparison to the feedforward networks of figures 2.3 and 2.4). As figure 2.5 shows, the network has an input vector **X** applied directly to the computational neural layer. The subsequent output is fed back to the distribution layer and then subsequently through the weight coefficients to produce a new output which is then fed back to the input layer and so on. The process ends when the network produces a stable, unchanging output vector.

To quantitatively describe the network of figure 2.5, it will be assumed that the network outputs are discrete-valued (0 or 1) and that the activation function for the neurons is that of a simple threshold function (this was the original model of the Hopfield network). Specifically, the outputs from the network obey the relations:

$$net_j = \sum_{\substack{i=1 \\ i \neq j}}^{n} w_{ij} out_i + x_j \qquad (2.18)$$

and

$$\begin{aligned} out_j &= 1 \quad ; \quad if\ net_j > T_j \\ out_j &= 0 \quad ; \quad if\ net_j < T_j \\ out_j &\ unchanged \quad if\ net_j = T_j \end{aligned} \qquad (2.19)$$

An invaluable contribution to the development of recurrent neural networks came in 1983 when Cohen and Grossberg [1983] proved that a recurrent network is stable whenever its interconnection weight matrix is symmetrical with a vanishing diagonal, i.e., $w_{ij} = w_{ji}$ and $w_{ii} = 0$ (the elements w_{ii} in figure 2.5 are indicated as set to zero and thus the sum in equation (2.18) for net_j involves only the outputs from all the other neurons). Defining a global, artificial neural network 'energy' (Liapunov) function (Simpson [1990]) as

$$E = -\frac{1}{2}\sum_i \sum_j w_{ij} out_i out_j - \sum_j X_j out_j + \sum_j T_j out_j \qquad (2.20)$$

it can be readily shown such a function monotonically decreases in value each time the network changes state (i.e., its input and output values). Ultimately this function reaches a minimum value which is unaffected by further iterations of the network, i.e., the network exhibits stability. It should be noted that the Cohen-Grossberg theorem is a

A Single-Layer Recurrent (Hopfield) Network

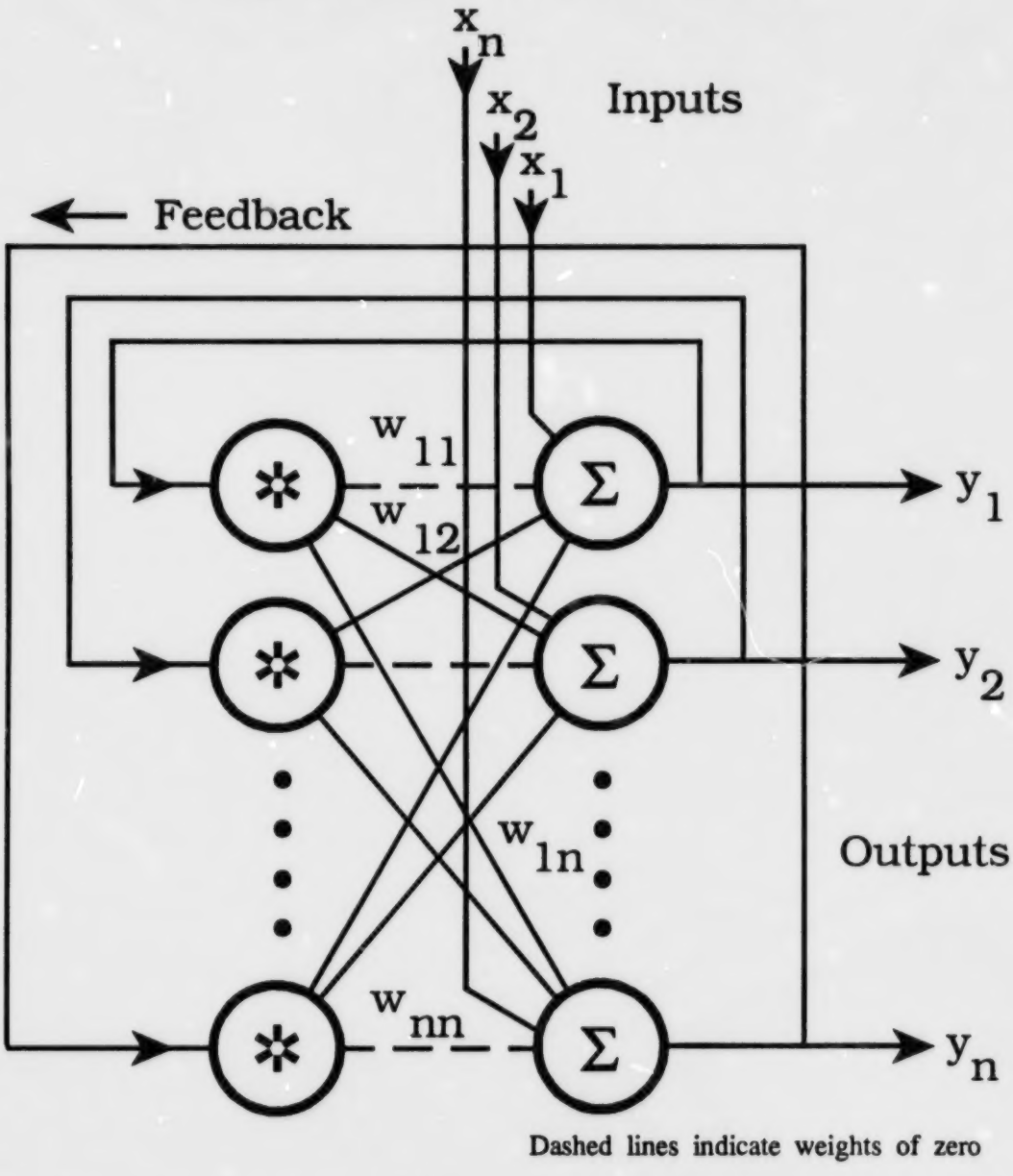


Figure 2.5 - The basic architecture of a single-layer Hopfield network

sufficient but not necessary condition; many networks are stable even though they do not satisfy the symmetry conditions stated above. The weights in the Hopfield network of figure 2.5 are determined, subject to the conditions $w_{ij} = w_{ji}$ and $w_{ii} = 0$, in a manner similar to that for the single-layer perceptron through a training session with known input-output vector pairs. More commonly, Hopfield-like networks are trained using statistical methods and these will be discussed in the following section.

Closely related to the structure and operation of Hopfield networks are a class of networks known as associative memories. The elementary Hopfield network of figure 2.5 is capable of operating as an associative memory by encoding the desired output vectors out_m, m = 1,2, ,M into the interconnection weights according to

$$\begin{aligned} w_{ij} &= \sum_{m=1}^{M} out_{i,m} out_{j,m} \quad ; \quad i \neq j \\ &= 0 \quad ; \quad i = j \end{aligned} \tag{2.21}$$

where $out_{j,m}$ is the jth component of the mth output vector. Now when an input vector is introduced into the network (when used as an associative memory, the inputs in figure 2.5 would be applied instead at the output nodes) the network will eventually produce one of the encoded output vectors which is closest (in the least-mean-square sense) to the input values. However, such a network is NOT operating simply as a "look-up" table. Rather, the neural network will produce an exact output vector even when the input vector is noisy, incomplete, or distorted relative to that used to train the network.

The relatively simple associative memory just described is of limited use since it represents an autoassociative memory, i.e., a "memory" or input vector can be corrected or completed to produce an exact original of itself. This limitation is the direct result of the single-layer structure of the network illustrated in figure 2.5.

The two-layer, bidirectional associative memory as illustrated in figure 2.6 is capable of heteroassociative behavior, i.e., it accepts an input vector at one set of neurons and eventually produces an associated or related output vector at the other set of neurons. An essential characteristics of networks such as the one shown in figure 2.6 is that the interconnect weight matrices for the two layers are related, specifically one matrix being the transpose of the other. The input and output vectors in this network are related as

$$\begin{aligned} Y' &= F(XW) \\ X' &= F(YW') \end{aligned} \tag{2.22}$$

where F is the neuron transfer function (most commonly the logistic function for such networks). The weight coefficients w_{ij} are encoded according to

Bidirectional Associative Memory (BAM)

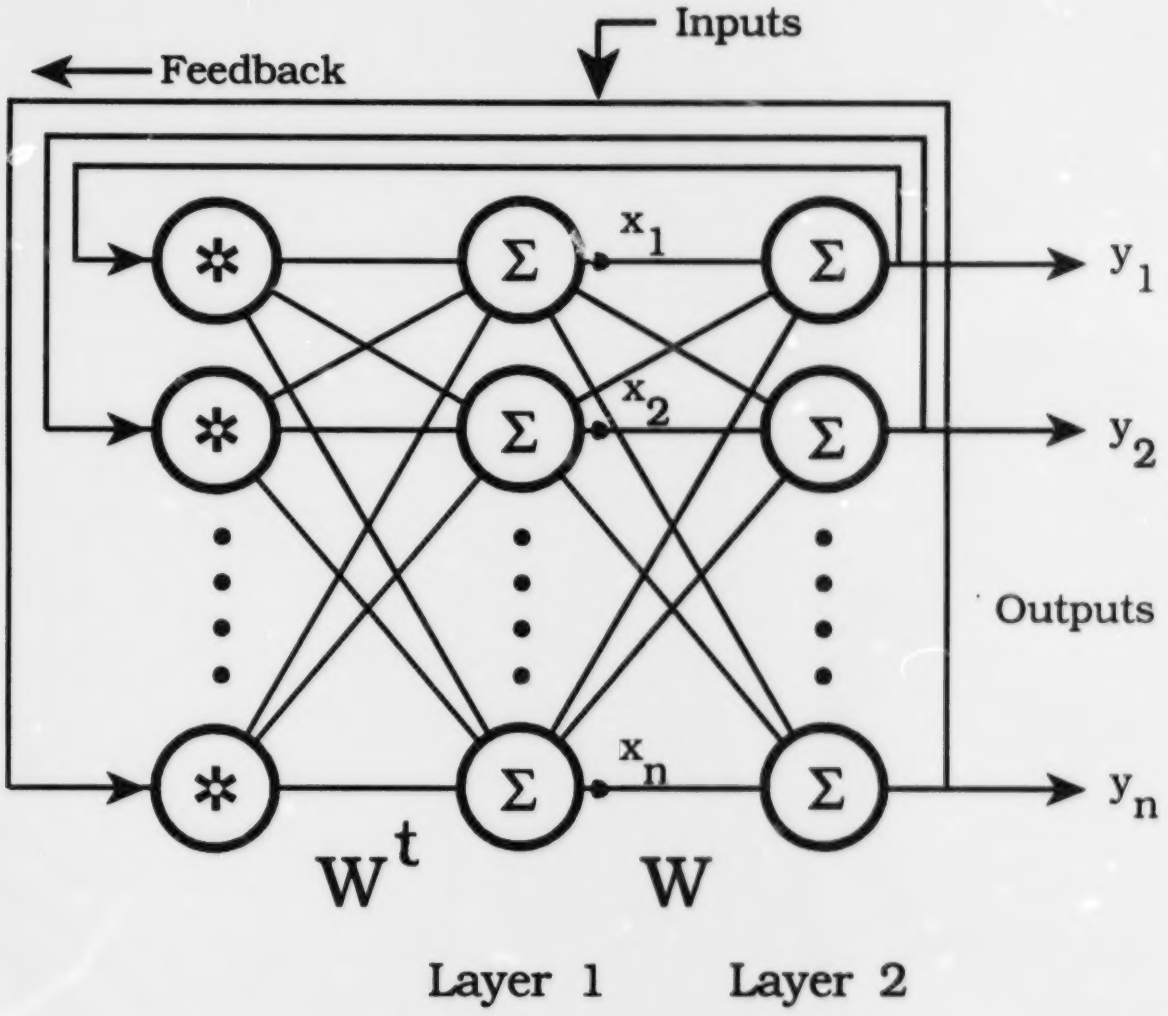


Figure 2.6 - The basic architecture of a bidirectional associative memory

$$W = \sum_{m=1}^{M} A_m^t B_m \qquad (2.23)$$

where $\{A_m, B_m\}$, m = 1,2, ,M are the (training) set of paired memory vectors with **A** corresponding to the inputs and **B** corresponding to the outputs. In operation, an input vector **X** is first applied to the network in figure 2.6 and the corresponding output **Y** is calculated via the weights **W**. This output is then fed through the weights $\mathbf{W}^t$ to produce a new value for **X** which then leads to a new value for **Y** and so on. Eventually the input and output vectors will stabilize to one of the training pairs $\{A_m, B_m\}$.

Bidirectional associative memories, which are capable of generalization, producing correct and exact outputs for corrupted or incomplete inputs, remain an object of much interest and research. Many questions remain to be answered completely, particularly difficulties in understanding the limitations to the memory capacity of such networks. In addition to the bidirectional associative memory example discussed above, other important variations on these networks have been advanced. Adaptive bidirectional associative memories, for example, have been studied which are capable of encoding an "idealized" output vector for a class of noisy or partial input training vectors. It should be noted that although the examples illustrated in this section have either implicitly or explicitly assumed discrete valued vectors and neurons exhibiting threshold behavior, all of these networks and associative memories have been generalized to operate with continuous valued vectors and transfer functions. The reader interested in further details of these and other forms of Hopfield and associative memory networks will find fairly complete descriptions given by Simpson [1990], Wasserman [1989], Khanna [1990], Miller [1987], Hopfield [1982, 1984], Lalanne, Taboury, and Chavel [1987], Kosko [1987], Tank and Hopfield [1987], and Dayhoff [1989].

2.6 Stochastic and Unsupervised Training Paradigms

Each of the neural network examples presented heretofore have been examples of neural systems trained by a supervised, deterministic paradigm. It is possible, however, to employ "unsupervised" learning along with statistical or stochastic methods for the adjustment of interconnection weight coefficients.

Supervised training involves adjustment of weight values as **pairs** of input/output vectors are used to train the network, the adjustment of weights always designed to reduce and minimize errors between calculated and desired output vectors. Unsupervised learning, in contrast, utilizes **only** input vectors in the training set and proceeds by monitoring the network response and adjusting the weight values to as to produce self-consistent output vectors and to group similar input vectors into classes which ultimately

will produce the same network response (in certain respects, the term "unsupervised" can be misleading and perhaps a term such as "directed learning" would be more precise). Several researchers are actively studying unsupervised learning paradigms, most notably Kohonen [1977, 1982, 1984, 1988] and his concepts concerning self-organizing neural networks and Grossberg and Carpenter [1986, 1987, 1988] with their extensive work on adaptive resonance theory. These studies are highly theoretical ones with the greatest emphasis placed upon the understanding of how various types of neural networks can function as an integrated learning system which exhibits, albeit modestly, characteristics analogous to biological neural systems.

A deterministic training algorithm is one for which the changes in the weight coefficients are determined solely by errors between training vectors and their calculated counterparts. Statistical training algorithms, in contrast, involve some form of pseudorandom changes to weight coefficients which are intended to optimize (usually to minimize) some global characteristic of the network. To illustrate the principles behind statistical training, a training algorithm known as Boltzmann training can be prescribed by the following iterative series of weight adjustments:

(1) Begin by defining a high value for an artificial network "temperature" T (see equations (2.24) and (2.25) below);
(2) Propagate an input vector $\mathbf{A}_r$ through the network to calculate the corresponding output vector $\mathbf{Y}_r$;
(3) Calculate the value of the network's objective function (a commonly used objective function is defined as the sum over all of the output neurons of the square of the differences between calculated ($\mathbf{Y}_r$) and expected ($\mathbf{B}_r$) output vector components);
(4) Introduce a random weight change and recalculate the network output along with the change in the objective function resulting from the weight change;
(5) If the objective function is improved (reduced for the definition described in (3) above), retain the change;
(6) If the objective function is worsened by the weight change in (4), calculate the probability of accepting that change from the Boltzmann distribution as:

$$P(c) = e^{-\frac{c}{kT}} \qquad (2.24)$$

where

$P(c)$ = *probability of a change of c in the objective function*
k = *a constant analogous to Boltzmann's constant* (2.25)
T = *the artificial network temperature*

(7) Choose a random number "r" from [0,1] and, if P(c) > r, retain the weight change or, if P(c) ≤ r, return the weight to its previous value;
(8) Repeat steps (4) to (7) for EACH weight in the network, reducing the value of T until it is acceptably low; and
(9) Return to step (2) and select a new training vector pair until the network has been trained on all the training vectors (perhaps repeatedly), stopping only when the objective function is acceptable for each of the training pairs.

This approach to network training is highly analogous to high temperature annealing of metals and thus if often referred to as "simulated annealing". The Boltzmann "machine" has been extensively studied and combined with other training concepts by several workers. Geman and Geman [1984] have shown that rate of the reduction of the network's "temperature" must follow

$$T(t) = \frac{T_o}{\log(1 + t)} \tag{2.26}$$

if convergence to a global minimum of the objective function is to be achieved (T_o is the initial temperature and "t" is an artificial time). This result predicts very slow convergence and long computation times, a prediction borne out in practice where Boltzmann "machines" frequently require unrealistically long training times.

To overcome some of the convergence difficulties of the Boltzmann algorithm, other (physical) models have been adopted for the probability distributions. One of the most successful has been Cauchy training which is based upon the distribution

$$P(c) = \frac{T}{T^2 + c^2} \tag{2.27}$$

which results in a temperature reduction rate as

$$T(t) = \frac{T_o}{1 + t} \tag{2.28}$$

and, accordingly, drastically reduced training times. Network paralysis, however, is a problem which is particularly severe for the Cauchy paradigm. Finally, Hinton and Sejnowski [1986] have demonstrated that it is possible to apply a statistical training algorithm such as the Boltzmann approach to a network already trained by a deterministic algorithm and achieve substantial improvements in the overall performance of the neural network.

2.7 Taxonomy of Artificial Neural Systems

One particularly insightful scheme for the categorization of artificial neural paradigms divides the paradigms into, (a), networks which are trained by either supervised or unsupervised methods and, (b), networks which, once fully trained, operate as either feedforward or recurrent networks. Figure 2.7 illustrates the taxonomic distinctions which result from such a choice of classification criteria (this figure is adapted from a similar one given by Simpson [1990]). Other characteristics such as deterministic versus statistical training methods or discrete versus analog calculations would serve poorly as a taxonomic criterion since many types of networks have been either operated in both of these modes and/or have utilized some combination of these characteristics simultaneously (e.g., the combined use of deterministic and statistical methods to train so-called Hopfield-Boltzman networks). Thus certain network types could conceivably be placed in two or more of the categorizations of figure 2.7 (Hopfield networks, for example, also exist as feedback recall networks which learn by supervised training algorithms). Instead, that figure shows any given network paradigm in that classification which has proven to be the most effective and studied for that particular network paradigm.

As was mentioned very briefly in the introduction to this chapter, the complete specification, independent of the method of implementation, of an artificial neural network involves a disquieting long list of choices, not all of which are quantifiable (here we restrict attention to self-contained, structured neural networks and omit the more esoteric and largely heterogeneous systems represented by concepts such as adaptive resonance theory and self-organizing maps). Included on such a list would be:

(1) The choice of artificial neuron and the specification of all parameters which describe the neuron activation function;
(2) The number of neuron layers (and possibly the type of artificial neuron in each) along with the number and distribution of the neurons in each layer;
(3) The precise geometry of the interconnections including any feedback paths along with the input and output nodes;

Taxonomy of ANS Paradigms

Operation →	Feedforward Recall	Feedback Recall
Encoding →		
Unsupervised Learning	• Counterpropagation (CPN) • Learning Matrix (LM) • Linear Associative Memory (LAM) • Optimal Linear Associative Memory (OLAM) • Learning Vector Quantization (LVQ) • Sparse Distributed Associative Memory (SDM) • Fuzzy Associative Memory (FAM) • Driver Reinforcement Learning (DRL)	• Adaptive Grossberg (AG) • Shunting Grossberg (SG) • Binary Adaptive Resonance Theory (ART1) • Analog Adaptive Resonance Theory (ART2) • Discrete Hopfield Net (DH) • Continuous Hopfield Net (CH) • Birectional Associative Memory (BAM) • Temporal Associative Memory (TAM) • Adaptive Birectional Associative Memory (ABAM)
Supervised Learning	• Perceptron • Adaline & Madaline • Backpropagation (BP) • Boltzmann Machine (BM) • Cauchy Machine (CM) • Adaptive Heuristic Critic (AHC) • Avalanche Matched Filter (AMF) • Associative Reward/Penalty (ARP)	• Fuzzy Cognitive Map (FCM) • Brain-State-in-a-Box (BSB)

Figure 2.7 - Taxonomy of artificial neural systems

(4) The choice of operating the network as a discrete- or analog-valued network;
(5) The detailed specification of the learning algorithm used to train the network (supervised, unsupervised, deterministic, statistical) including the criteria used to evaluate the training process and to determine when the training is completed; and
(6) The specification of the training set (either input vectors or a set of paired input/output vectors) used by the learning algorithm and any methodology employed for the ordering in the presentation and selection of the training data.

Note that several of these items such as (5) above would themselves normally entail a further sublist of choices to be made by the network designer (for example, the choice of backpropagation training leads to further choices involving modifications or "improvements" to the algorithm such as those described briefly in section 2.4).

It is this profusion of choices which has led to the creation of so many distinct and independent neural network paradigms coupled with the fact that there does NOT exist any "master" or generalized theory of neural networks which would encompass, at least in theory, all the various models (it is not clear, at present, that such a "master" theory may even exist). Thus this plethora of paradigms can be viewed in two extremes - one view would state that the diversity and multiplicity of paradigms reflects the richness and profound potential of this new field of research while the other view would, somewhat cynically perhaps, that this very diversity speaks poorly for the field, indicating a directionless, confused area of research. The truth, as is often the case, lies somewhere between these two extremes.

The present and future development, particularly the move towards ever larger and more powerful networks, depends more upon the acquisition of a general, systematic understanding of neural network characteristics than the development of novel technology or designs to implement the neural hardware. Of course, advances in the latter will almost certainly assist greatly the former; the point is that the technological future will be muted in the absence of substantive progress in the fundamental understanding of neural network operation. A great deal of progress in the field has been achieved in the past decade alone, much of it in an "empirical" fashion, i.e., questions of optimization, network size and structure, and global minimization of error have frequently been answered on a trial-and-error basis using realistic but highly parametric models. This trend seems certain to continue unabated.

The other factor which tends somewhat to mitigate the apparent proliferation of diverse paradigms shown in figure 2.7 (keeping in mind that this figure is far from exhaustive) is that many of the paradigms mentioned there are, to be forthright, of transient and relatively minor importance either because (a), that particular approach is extremely specialized and hence limited in applicability to real problems, (b) proves

ultimately to possess some 'fatal flaw' in its characteristics, and/or, (c), is superseded by a paradigm which is more general, efficient, and dependable.

Each of the paradigms identified in figure 2.7 possesses characteristic strengths and weaknesses with respect to such important properties as computational complexity, training times, susceptibility to network paralysis, entrapment in local minima, ability to generalize, memory capacity, adaptability, and suitability to hardware implementation(s). The discussion of the representative paradigms given in this chapter has, of necessity, been limited to several specific elements. Comprehensive and detailed treatments of artificial neural network taxonomy including detailed discussions on a paradigm-by-paradigm basis may be found in Simpson [1990], DARPA [1988], Khanna [1990], Wasserman [1989], Miller [1987], and Dayhoff [1989].

2.8 Implementation of Artificial Neural Systems

There exists three largely distinct methods for the implementation of neural networks, namely as simulations, electronic, and optical (including electro-optical) implementations as illustrated in figure 2.8. To date and for the foreseeable future, the preponderance of the work and results presented in neural networks is carried either largely or exclusively via simulation of neural nets on computers. These simulations are either purely software-based or, more commonly, utilize dedicated coprocessor hardware (noting that the latter is, more accurately, a true electronic implementation as indicated in figure 2.8).

There are numerous reasons why software/hardware simulation of neural networks is so predominant in this field of research:

(1) The veritable ubiquitousness of conventional computers coupled with readily available, inexpensive commercial software/hardware for network implementations has made it possible for scores of individuals and research groups to actively participate in and make contributions to the field with very few if any costly investments;

(2) As discussed in the previous section, the principal challenges to the short term progress in neural network research consist of solving many problems concerning the more fundamental characteristics of neural systems as systems, i.e., to gain further understanding and command of such aspects as optimization of architectures and training paradigms;

(3) Related to the point immediately above is the versatility of the simulation approach, i.e., it is straightforward to utilize a variety of radically different architectures (models), neuron models, and learning algorithms via simulation whereas such versatility is pragmatically prohibited for strict hardware implementations;

(4) With very few exceptions, full hardware implementations of neural networks using either electronic and/or optical technologies are invariably

Figure 2.8 - Methods of implementing artificial neural networks

constructed with fixed interconnect weights, i.e., the training of such networks is done off-line using software/hardware simulators since this vastly simplifies the hardware necessary to the implementation; and

(5) Even (or, rather, especially) when the object of the research is to hardwire a neural network, the simulator begins as essential accessory by which alternate designs and variations can be quickly and realistically tested and evaluated.

The principal drawback to the use of simulators is their relative inefficiency and slow speed, particularly for large or computationally intensive paradigms. Dedicated electronic processors such as the Delta II and Anza Plus accelerator cards offer substantial increases in neural computational power. For even larger models or more complex neural systems, the use of supercomputers or the newer "parallel" computers becomes pragmatically necessary.

The dedicated attached processors such as the Anza Plus and dedicated buss-orientated processors such as TRW's Mark series represent electronic implementations of artificial neural systems which use conventional electronic components and software which essentially maps the simulation onto electronic chips. These represent virtual electronic implementations in that any given electronic processor may be use, via time multiplexing, to act as several artificial neurons. A fully-implemented electronic implementation, on the other hand, would have separate hardware circuitry for each elementary neural processor. While the latter will offer the ultimate in speed for neural computations, the former makes much more efficient and versatile use of present, "overly powerful" single-chip microprocessors. Nevertheless, the development of fully-implemented electronic is being pursued at the research level by several research groups. In addition, work on the development of VLSI neural "chips" has increased dramatically in the past several years and several commercial laboratories are developing commercial prototypes of such circuitry. In general, these circuits are relatively modest in terms of numbers of neurons (the largest to date has 256 neurons in total) but all expectations point to a rapid increase in this number within the near future. Treleaven, Pacheco, and Vellasco [1989] and Mead [1987] give detailed reviews of VLSI implementations of neural circuitry, Verleysen and Jespers [1989] describe an analog VLSI implementation of the Hopfield net, while Simpson [1990] and Miller [1987] give more general reviews of all types of electronic implementations for a variety of network paradigms.

Although generally conceded to hold the greatest long-term potential for large scale neural network implementation, optical implementations (including hybrid optical/electronic and opto-electronic) are both the fewest and least developed of the three methods depicted in figure 2.8. Since these developments will be discussed in the next chapter, the topic of optical neural network implementations will not be pursued here beyond referring the interested reader to general reviews offered by Caulfield, Kinser, and Rogers [1989], Anderson [1989], Miller [1987], Athale [1989], Athale and Davis [1988], Wagner, Giles, and Psaltis [1990], Psaltis et al [1989], and Wasserman [1989].

3.0 OPTICAL NEURAL NETWORKS

3.1 Introduction

As depicted in figure 3.1, the field of optical neural networks is symbiotically related to the fields of classical neural networks, classical optical computers (including recent and ongoing development of digital optical computing technology), and optical device research. Indeed, at present, this field could more aptly be termed the optical implementation of classical neural network paradigms using classical optical signal processing technology. Only a relatively small number of neural network paradigms have been implemented optically to date and most of these have involved the use of electro-optic devices. In addition, although some of these implementations are optically complex, the networks they represent are small and prototypical in nature when compared to those implemented via simulation and/or electronically.

Despite these observations, these exists a broad consensus in the neural network community that optics MUST play an essential role in the development of neural network technology, particularly as it progresses towards large scale, more powerful networks. The reasons underlying this belief stem from the many advantages inherent in the use of (laser) light as a communications and/or processing medium. The most important of these advantages are listed in figure 3.2.

Since one of the most fundamental and axiomatic characteristics of artificial neural networks is the dense interconnectivity of the processing elements, it is fair to say that the first three attributes listed in figure 3.2, namely the intrinsic parallelism of light along with its capability for free space, directed interconnections in two- and three-dimensional configuration, are of principal importance in the implementation of neural networks by optical means. In contrast, the limitations experienced in all-electronic technology concerning interconnectivity in what are essentially two-dimensional structures are formidable and well-known. Large fan-in and fan-out is impractical for VLSI technology since each connection requires a unique, discrete channel with sufficient electrical isolation. The penalties in terms of the power required and the additional space needed to implement large fan-in/fan-out values for integrated electronics are prohibitive. In addition, the constraints on electrical interconnection of distinct electronic layers (either intrachip or chip-to-chip) make this plausible only for small numbers of processing elements/layer with an attendant small number of total interconnects (indeed, the use of light as a layer-to-layer and device-to-device communications medium is one of the first and most vigorously studied areas of optical interconnect research).

These remarks are not to be interpreted as stating that optics is **THE** solution to the problem of neural network implementation. Rather, optics offers **A** solution to this problem. One way to illustrate the domains for which optical implementations of neural networks become desirable (arguably essential) is to consider the relationship between the number of elementary processors (neurons) and the number of interconnects for a

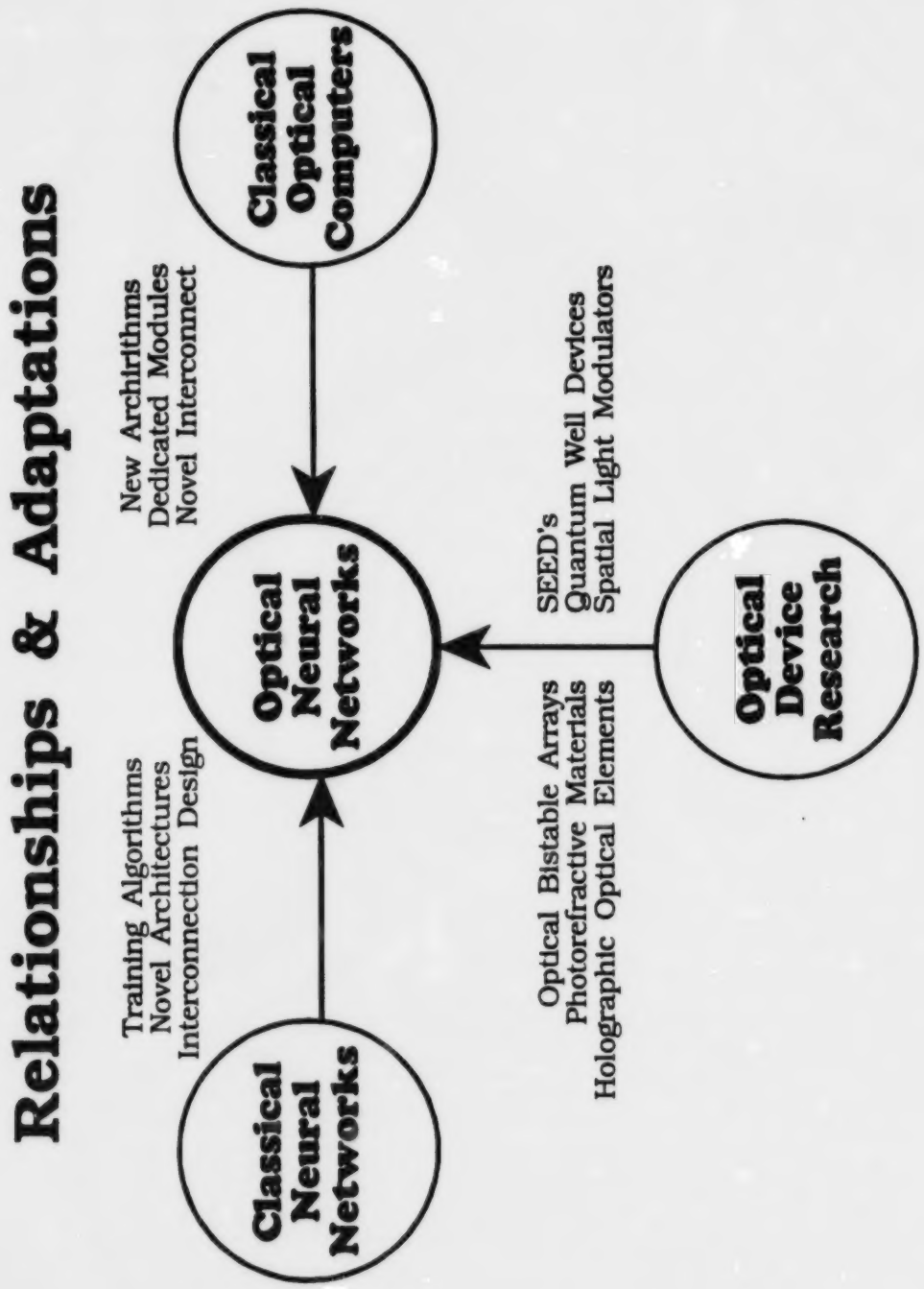


Figure 3.1 - Optical neural network relationships

Advantages of Optical Signals

- Inherent, intrinsic parallelism in capability to process and communicate with light
- Ability to process and transport optical signals in two- and three-dimensional geometries
- Free space, independent, directed optical interconnections
- High space- and time-bandwidths allow for many simultaneous optical channels
- Realistic potential for sub-picosecond processing, with switching times limited by device size and optical response characteristics
- Complete immunity from electromagnetic interference enables extremely high packing densities for devices and interconnects
- Flexibility to use different optical parameters as the information carrying variable: intensity, amplitude, phase, polarization state, and spatial position
- Optical signals can exploit time-, wavelength-, and polarization-multiplexing simultaneously
- Multi-valued logic can be readily implemented by the use of a many-valued variable such as spatial position or polarization

Figure 3.2 - Intrinsic advantages in the use of optical signals

prototypical neural network consisting of two layers, each with N neurons. If each neuron in the first layer is connected to only one neuron in the second, a total of N interconnections results. Any fewer would be described as "underconnected". In contrast, if each neuron in the first layer is connected to every neuron in the second, then there would be a total of N^2 interconnects. Any more than this would constitute an "overconnected" (or redundantly connected) network. Figure 3.3 (adapted from Caulfield, Kinser, and Rogers [1989]) illustrates the domains of optical and electronic technologies as functions of the number of elementary processors versus the number of interconnects. The two lines in that figure correspond to the minimally connected and fully connected boundaries noted above. At present, even the most ambitious neural networks under study would lie well down on the lower left corner of that chart, i.e., well inside the 'electronics domain' (in terms of brains, the best hardwired neural networks at present correspond approximately to the common earthworm with its 10^3 neurons). Future networks with interconnections numbering 10^6 and greater will probably belong exclusively to the optical domain. In those regions where the two technologies overlap, optical implementations must seek other advantages in characteristics such as power consumption or compactness. Figure 3.4 offers explicit albeit approximate guidelines to the delineation of the electronic and optical domains in terms of numbers of interconnects and neurons and in terms of the basic computational power measure for neural networks, interconnects/second.

For taxonomic purposes, optical implementations of neural networks may be divided into two broad classes. One class uses the variable transmittance of some two-dimensional device (examples include spatial light modulators, photographic film, and liquid crystal light valves) as the means of implementing the interconnection weights while the other category uses the (programmable) diffraction efficiency of a holographic grating for the same purposes (the former group would include the use of two-dimensional arrays of nonlinear optical devices such as bistable etalons or SEED's). Both groups make extensive use of optical matrix vector multipliers and other "classical" optical computing devices as the means for spatially broadcasting and collecting the weighted outputs from sources and detectors. The next two sections will present archetypical examples from each of these two categories along with discussions of their relative strengths and weaknesses. A final section will briefly discuss future developments of optically implemented neural networks and some of the key issues and problems faced by this research direction.

3.2 Optical Associative Memories and Hopfield Networks

A fundamental optical configuration which is at the heart of many optical implementations of neural networks and associative memory networks is the matrix-vector multiplier illustrated schematically in figure 3.5. An array of light emitting diodes acts as the source or inputs. The light from each diode is spatially distributed in one direction by the adjacent cylindrical lens - in the figure, each diode illuminates a horizontal row of five pixels in the spatial light modulator. The light passes through the individual elements of the spatial light modulator, each pixel transmitting "w_{ij}" of the incident light, i.e., the

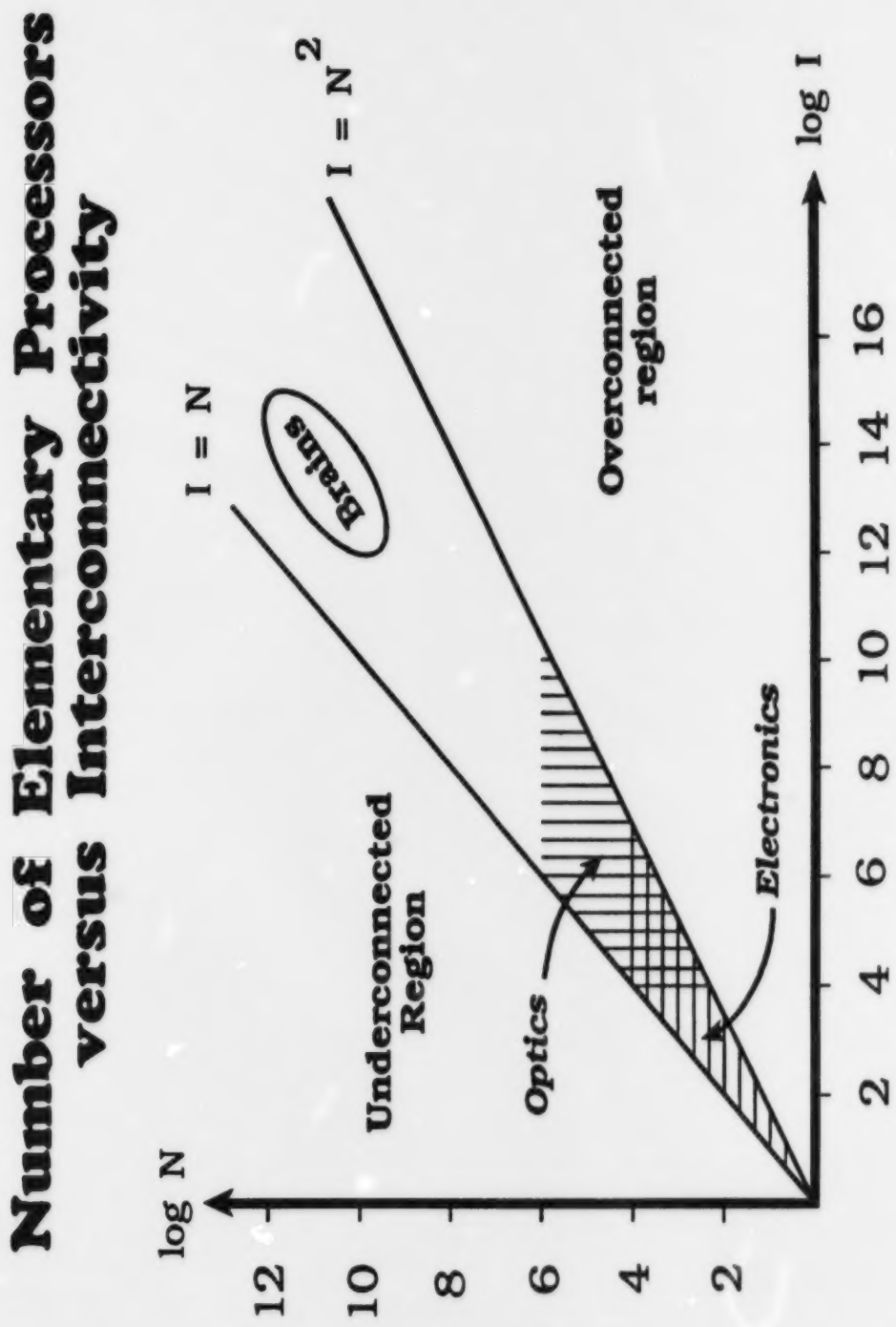


Figure 3.3 - Technological domains as functions of numbers of processors and interconnects

Electronic & Optical Domains

Characteristic	Electronic	Optical
Local Interconnects	✔	
Global Interconnects		✔
10^{1} - 10^{5} Interconnects	✔	
10^{6} - 10^{12} Interconnects		✔
10^{4} - 10^{9} Interconnects/sec	✔	
10^{10} - 10^{20} Interconnects/sec		✔
10^{1} - 10^{4} Neurons	✔	
10^{4} - 10^{8} Neurons		✔

Figure 3.4 - Categorization of electronic and optical domains

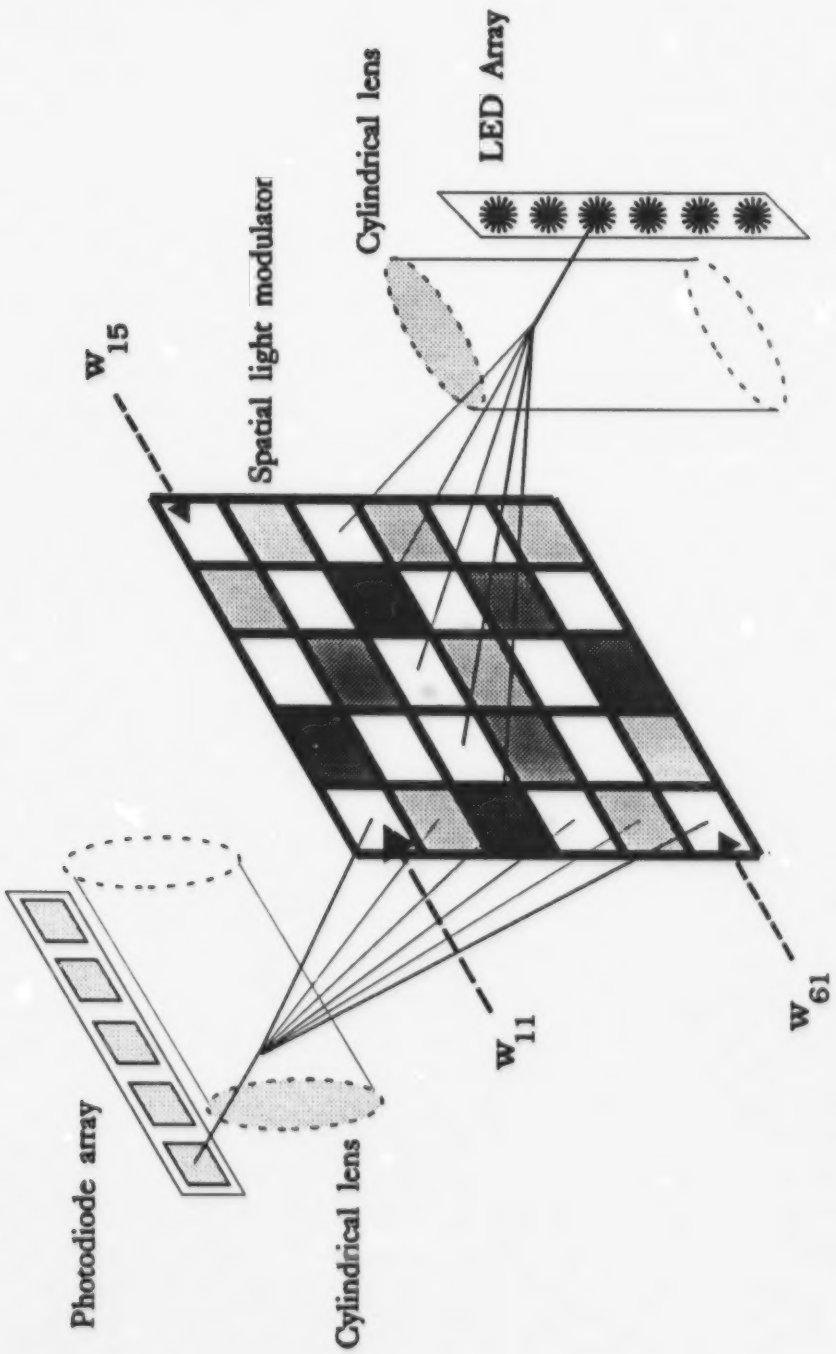


Figure 3.5 - An optical matrix-vector multiplier

transmittances serve as the synaptic weights. The transmitted light is then collected by a second cylindrical lens which is at right angles to the first. This second lens collects (i.e., spatially integrates) and focuses to each of the photodiodes in the adjacent array all of the light from one vertical column of six pixels in the spatial light modulator. Although not shown in figure 3.5, in practice two other, large aperture cylindrical lens would normally be placed close to and on either side of the spatial light modulator and would serve to collimate the light entering and leaving this device (this substantially reduces the loss of light from LED to photodiode).

This relatively simple optical configuration in fact performs the matrix-vector multiplication given in equation (2.1) where the "x_i" would represent the LED intensities and the "net_j" would be the resultant output signals from each of the photodiodes. Indeed, the "equivalent" neural circuit diagram of figure 3.5 would be that of a single-layer perceptron, figure 2.3, with six inputs and five neurons. In figure 3.5, the first cylindrical lens plays the role of the distributive layer, the weight coefficients are just the pixel transmittances of the spatial light modulator, and the photodiode array itself acts as the neuron layer by summing these weighted, optical signals.

Upon the addition of an appropriate (electronic) feedback loop from the photodiode array back to the LED array and the use of a square spatial light modulator, this speciously simple optical circuit becomes a optical Hopfield neural network (also an autoassociative memory) such as that illustrated by figure 2.5. To satisfy the stability requirements, the spatial light modulator transmittances would be symmetric with zero transmittances along the matrix diagonal. The electronic circuitry from the photodiode array to the LED array would provide, in addition to the necessary feedback, thresholding (the activation function) and gain.

The matrix-vector multiplication carried out by the optical network of figure 3.5 is performed in parallel. With suitable high-speed photodiodes and LED's, the entire matrix-vector multiplication can be performed on a nanosecond time scale. In addition, the speed capability is almost entirely independent of the size of the array so that the network can be scaled upwards significantly without increasing the time required for the computation. Although the optical Hopfield network can be operated with a fixed, analog weight mask such as a photographic film, the use of dynamic elements such as a spatial light modulator or liquid crystal light valve enables the network to operate as an adaptive network including the capability for on-line training. At present, technological limitations to spatial light modulators and liquid crystal light values pragmatically discourage their use as analog devices. They are, instead, operated as binary weights but this shortcoming is expected to be overcome in the near future.

The reader may recall from chapter 2 that the transition from an autoassociative memory, figure 2.5, to the far more useful heteroassociative or bidirectional associative memory, figure 2.6, requires the addition of a second neural layer. By cascading two of the optical networks of figure 3.5, using the output from the first photodiode array to drive the LED array of the second stage, an optical bidirectional associative memory can

be implemented. A much more elegant and efficient method of achieving this was demonstrated by Kosko [1987]. Exactly the same optical configuration of figure 3.5 was employed except that the LED and photodiode arrays were replaced with detector-source arrays (paired LED-photodiode arrays with each LED-photodiode pair electronically coupled with gain and thresholding functions). In this manner, light from each array was transmitted to the photodiodes of the other which in turn drove their adjacent LED's which then fed back to the original detector-source array and so. The light signals thus propagated simultaneously in both directions, intersecting but not interacting. Ingeniously, the light traversing in one direction sees a weight matrix which is the transpose of the same weight matrix as seen light traversing in the opposite direction. This compact, simple, and elegant implementation of an optical bidirectional associative memory utilizing but one modulator device is an excellent example of the versatility and creativity often possible in the implementation of neural networks by optical means.

Optical implementations of Hopfield networks and associative memories using two-dimensional transmittance devices have been a source of considerable interest and research. Many variations and adaptations have been reported. For further details, the reader is referred to Farhat, Psaltis, Prata, and Paek [1985], Hopfield [1982, 1984, 1987], Jang, Jung, Lee, and Shin [1988], Psaltis and Farhat [1984, 1985], Athale, Friedlander, and Kushner [1986], Athale, Szu, and Friedlander [1986], Gindi, Gmitro, and Parthasarathy [1987], Jang, Shin, and Lee [1988], Johnson, Kranzdorf, Bigner, and Zhang [1989], Kinser, Caulfield, and Shamir [1988], and Kosko [1987].

3.3 Optical Neural Networks using Holographic Components

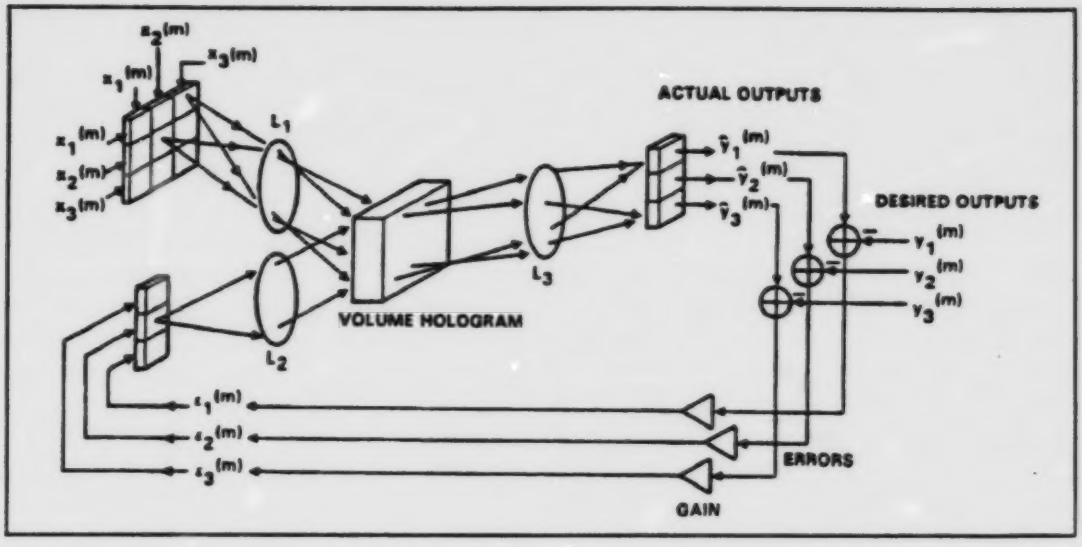
The examples discussed in the previous section centered around two-dimensional optical implementations in which the interconnection weights were represented by the pixel transmittances (or possibly reflectances) of either passive or electro-optic devices such as photographic film and spatial light modulators (nonlinear optical devices such as SEED arrays and bistable etalon arrays can also serve this function in two-dimensional configurations). The use of volume holographs, particularly in photorefractive index materials, to provide the weights via diffraction angles represents an attempt to fully exploit the three-dimensional nature of light. A volume or "thick" hologram differs from the better known planar hologram in that a volume hologram stores the interference pattern comprising the image or pattern throughout the three-dimensional volume available to it. The hologram's reference beam is a planar wave and represents a summation of waves, each emanating from a different point source in the reference plane. Each of these 'point source' reference waves intersects the hologram during exposure at a slightly different angle from the other reference waves. In the volume hologram, because of the optical interaction in the third dimension (along the thickness of the hologram), the hologram is able to resolve these small differences in the angle of incidence of the reference beam and, upon illumination, can selectively reconstruct any of a series of two-dimensional images or patterns stored in the holographic recording. This is a brief and somewhat simplified explanation of volume holograms; the reader

wishing to learn about these optical elements in greater detail may refer to Cohen [1986], Dunning, Marom, Owechko, and Soffer [1985], Owechko, Dunning, Marom, and Soffer [1987], and Psaltis, Brady, and Wagner [1988].

The lower illustration in figure 3.6 (from Psaltis, Park, and Hong [1988]) shows schematically how a volume hologram can be used as the interconnection medium for arrays of sources (inputs) and/or detectors (outputs) in at least two distinct ways, i.e., mapping a two-dimensional array onto a vector array and vice versa (the first is referred to as $N^2 \rightarrow N$ mapping, the latter as $N \rightarrow N^2$ mapping). Such volume holograms when used as interconnection distribution devices are frequently placed in an optical resonator or ring cavity in order to provide continuous optical feedback. A variety of optical neural networks and associative memories have been built or proposed using volume holograms (see Anderson [1987], Peterson, Redfield, Keeler, and Hartman [1990], Psaltis, Brady, and Wagner [1988], Wilde and Hesselink [1989], Xu, Yuan, Yu, Xu, and Xu [1990], Dunning, Marom, Owechko, and Soffer [1985], Owechko, Dunning, Marom, and Soffer [1987], and Shamir, Caulfield, and Johnson [1989].

If passive holograms (usually photographic film) are used in optical implementations of neural networks, then such networks must necessarily be non-adaptive since the weight coefficients are a priori fixed. Much excitement presently exists over the use of various photorefractive index materials as the holographic medium since laser light is used dynamically to write and read the holographic grating. Such an ability allows for the design of neural networks and associative memories which are adaptive once trained and which could be trained on-line. The schematic on one such design for an optical associative memory is shown at the top of figure 3.6. Here the corrections (i.e., the equivalent of the δ factors of chapter 2) to the weight coefficients (the diffraction angles of the hologram) would be fed back via processing electronics which would control an array of LED or laser sources. These in turn would rewrite elements of the grating within the volume hologram and thus provide the weight adjustments. Such a configuration would allow for very rapid network training since all of the arrays are processed in parallel (the limiting factor would be the electronic feedback loop).

The use of holograms, either planar or volume, in optical neural networks is particularly suitable for the construction of optical associative memories. A singularly explicit example of such an application was reported by Abu-Mostafa and Psaltis [1987] and summarized by figure 3.7. Images were stored in two (planar) holograms within an optical ring circuit containing a threshold device which has a reflectively on one surface which increases according to the brightness of the light impinging upon the other surface. The light from the input image was reflected from this device and directed back through one of the holograms containing four images, creating output patterns of light with a brightness dependent upon how well the input and stored images match. The lens/pinhole combination spatially separates the four stored images. This light then passes through the second hologram so that the brightest image of the four stored there corresponds to the best match present in the first. The brighter light thus impinges upon the threshold device and thus the best match to the input images from the stored images



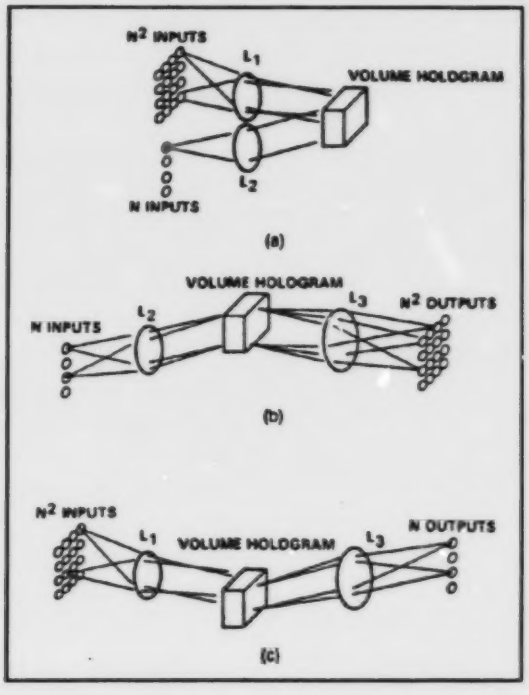
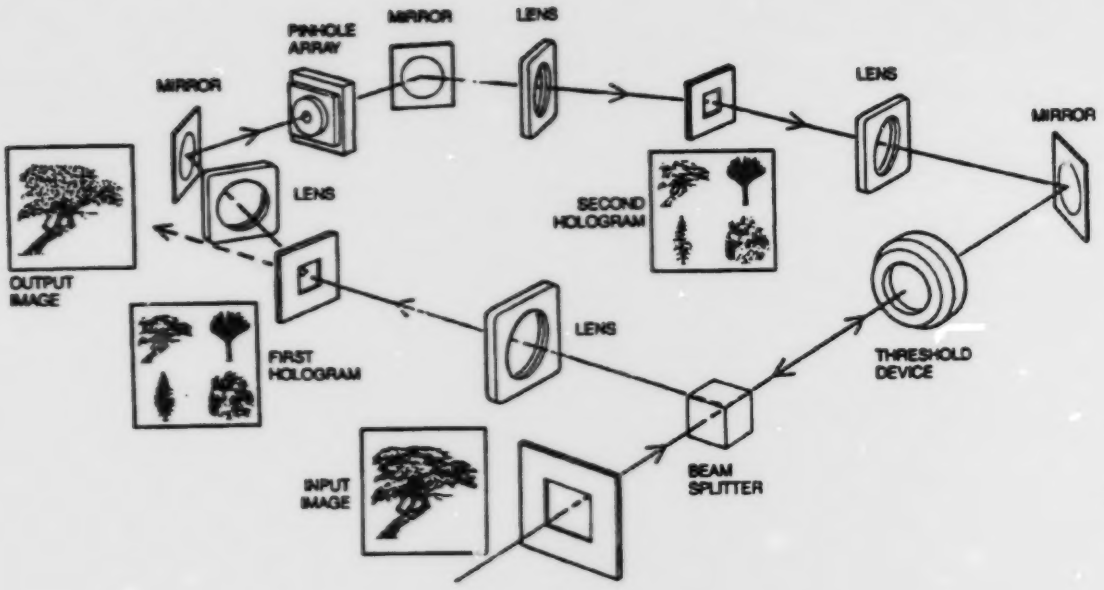


Figure 3.6 - The use of volume holographic elements in optical neural networks and associative memories

(From D. Psaltis, C. H. Park, and J. Hong, "Higher Order Associative Memories and Their Optical Implementations", Neural Networks 1, 1988, pp. 149-163)



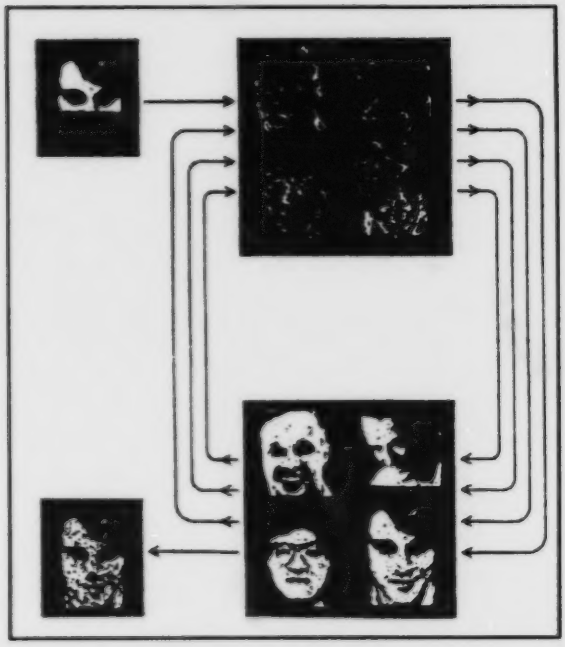


Figure 3.7 - An optical neural autoassociative memory. This device is capable of recalling a complete image when only a partial image is input.

(From Y. S. Abu-Mostafa and D. Psaltis, "Optical Neural Computers", Sci. Amer. 256 (March), 1987, pp. 88-95)

is preferentially reflected back through to the first hologram. Successive passes around the optical loop continuously reinforce this selection until the output image is that corresponding to the 'best fit' to the input. The lower illustration in figure 3.7 shows the result for a group of images of different faces. When only a part (approximately 1/2) of one of these images is input to the optical loop, the output image is an exact match to the complete stored image corresponding to the input. As limited as this network is, it behaves in a manner very similar to that by which humans perform the same function.

3.4 Future Research into Optical Neural Networks

The versatility of optics makes it possible to implement functionally equivalent neural networks in such a diversity of ways that it can be confusing to attempt to classify published results which, at first glance, appear to be completely unrelated one to the other. In some instances, clearly one configuration excels in achieving compactness and efficiency in its use of components and optical signals (the bidirectional associative memory of Kosko [1987] is one such example). Research interest remains high in the optical neural network community to find practical yet powerful optical implementations of classical neural network paradigms.

In the longer term, the approach based upon the use of volume holographic elements, particularly dynamic ones, offers the greatest potential. Theoretical estimates of the storage capability of photorefractive index materials place it from 10^8 to 10^{10} interconnects per cubic centimeter. The intrinsic three-dimensional capability of volume holograms make them the most general possible devices for different interconnection topologies. However, one major problem with photorefractive index volume holograms stems from the fact that any reading of the holographic grating also perturbs and changes it. Thus the grating eventually "unlearns" with use. Current research into these materials is concentrating upon methods to circumvent this fundamental problem (see the review by Wagner, Giles, and Psaltis [1990]).

In the shorter term, the two-dimensional approach using commercially available, "off-the-shelf" devices such as spatial light modulators remains extremely attractive since, (a), it has already been proven in a number of configurations and, (b), substantial improvements are expected in the size, range, speed, and accuracy of the devices currently used for the weighted interconnections. For example, recent developments in ferroelectric liquid crystal light valves have produced devices of 1000 x 1000 pixels with microsecond response times and sub-picojoule/pixel switching energies (Wagner, Giles, and Psaltis [1990] and Johnson, Kranzdorf, Bigner, and Zhang [1989]). Furthermore, as mentioned briefly previously, the two-dimensional implementations lend themselves readily to the adoption of nonlinear optical devices such as SEED arrays and bistable etalon arrays, both of which are presently areas of intensive research.

Figure 3.8 presents a summary of the types of optical implementations of neural networks ranging from all-optical to hybrid approaches in which light is used as the

Future Directions for Optical Neural Networks

Three concurrent approaches to technological implementation:

- Electronic processing technology with optical interconnections
- Opto-electronic/integrated optical processing with optical interconnects
- All-optical networks employing nonlinear optical devices as processors

There exists two principal difficulties to long-term progress:

- A more thorough understanding of neural network architectures and paradigms and their applicability to optical implementations
- Optical devices : improvements in existing devices and the development of novel, scalable, and more efficient devices

Figure 3.8 - Research directions for optical neural networks

communications medium and electronics plays a greater role in the processing tasks. To date and for the foreseeable future, those approaches involving degrees of opto-electronic and/or electronic processing will dominate over the all-optical approach. In addition, as noted in this figure, there are two principal problems which presently face the future development of optical neural networks. They are, (a), the further development and understanding in neural network architectures and learning paradigms and, in particular, the question of applicability of such paradigms to optical implementations, and, (b), the concurrent development of novel and improvements to existing linear and nonlinear optical devices. The first of these limitations is common to the field of neural networks in general and thus optical neural network development can expect to gain considerably from the research efforts of the much larger neural network community as a whole. The problems of translating neural network paradigms into optical implementations along with the development of improved optical devices, however, will remain the "exclusive" domain of optical researchers in this field. The work and reported successes to date have been very encouraging.

4.0 CONCLUSIONS AND RECOMMENDATIONS

4.1 Conclusions

- Neural networks, irrespective of the technology used to implement them, represent a genuinely novel and significant approach to building computational and information processing structures. The science and technology of neural networks, while still immature, have progressed substantially from their original development in the 1940's and 1950's. To date, many researchers have put forward significant demonstrations of neural network capabilities in information processing tasks such as speech recognition, pattern classification, robotics, and vision-related processing tasks. Even if future advances were discounted, neural network technology has already convincingly demonstrated that, in many applications, performance at least as good as (and frequently much better than) that from conventional computer technology is obtainable with neural network circuitry which is simpler, more compact, more robust, and easier to train and use in practice.

- At present, the single greatest impediment to substantive progress in neural networks including their applications is the lack of a comprehensive, general theory of networks which would point the way to the development of large scale nets as well assist in the delineation and categorization of the many, piecewise 'solutions' now existent as distinct paradigms. At the same time, there is a strong need for more powerful and more widely applicable training algorithms and novel computational models which could lead to both electronic and optical implementations. Directly related to this problem is the need for developing substantially better hardware capabilities in order to, in turn, directly improve the simulation and implementation capabilities for neural networks and associative memories.

- Neural networks appear to hold particular applicability to real-time signal processing and pattern recognition problems. Work in areas such as robotics, machine vision and other vision-related tasks, speech recognition, pattern classification, and data fusion should be strongly fostered. Commercial applications in such areas using neural network implementations have just begun to appear and will their numbers will increase rapidly by the end of this century.

- The speed, parallelism, ease of implementation, ability to learn and adapt, fault tolerance, and ability to generalize are characteristics of neural networks which make the technology extremely attractive in many military applications where robust, compact, computational devices are needed to process large amounts of sensorial information on a realtime basis.

- The importance of the role played by optics in future neural network implementations will increase in the future. It seems quite likely that the first truly

large scale and powerful neural networks will be implemented using optical signals and electro-optical components for many of the inter-processing tasks.

- The future will witness an increasingly close relationship between optical neural network development and optical communications systems in general. The ever increasing consolidation of communications and information processing concepts into unified 'intelligent' communications systems indicates that new found abilities to process information using optical technology may first find application within the field of optical communications.

4.2 Recommendations

The fields of neural network research and, in particular, optical neural network research, are changing very rapidly as new architectures along with advances in materials and devices are reported almost monthly. The vast preponderance of work done in these fields consists either solely or largely of simulative work with or without hardware support. While this situation will certainly remain true for some time to come, it is encouraging to note that more and more hardware implementations, particularly optically-based ones, are appearing in the published literature.

The intent of these recommendations relates to a long term research program conducted under the guidance of the Directorate of Optical Communications and Electrophotonics Technology (DOCT) at the Communications Research Center. The program would consist of an extramural program, chiefly contracting of research projects to Canadian universities and industries, conducted concurrently with an in-house (intramural) program. Since at the present time, DOCT has no direct experience in neural network research, the proposals given below are intended in part to enable DOCT to "catch up" in this area of research in order to pursue original and contributive work in the field.

(1) The acquisition of simulative capabilities for neural network architectures, design, and evaluation is a mandatory prerequisite to the development of any form of hardware implementation. A project to establish simulative capabilities would be the seminal project and, in addition to providing the necessary support for the eventual hardware development, would serve both as an educational and ongoing research tool in neural network development. The preferred method of implementing simulative ability is a dedicated processor board (such as the Delta II board from SAIC) since this gives substantially improved computational throughput in comparison to software alone. Such hardware is expensive, so that the alternative of a purely software driven neural network simulator must not be ruled out.

(2) Either concurrently or shortly after the initiation of simulative studies of neural and optical neural designs, work would begin on the experimental construction of

two-dimensional neural networks using spatial light modulators as the weight interconnect devices. Elementary autoassociators, optical Hopfield networks, and optical bidirectional associative memories are well understood and documented optical neural networks. In order to gain necessary "hands-on" experience, it is recommended that the initial stages of this work attempt to build such networks in order to study their characteristics and behavior. This would then lead DOCT to the point where a working knowledge of the optical hardware and elementary architectures would be in hand and allow original and novel research to go forward from that point. The decision to explore two-dimensional approaches to optical neural networks also brings the advantage that nonlinear devices such as may result from, for example, the continuing thin film (Langmuir-Blodgett) work within DOCT may be immediately applicable to the neural network research (conversely, established optical neural networks could, conceivably, be employed as a means of evaluating and characterizing novel, nonlinear optical devices). Finally, the two-dimensional neural network would lend itself to the incorporation of other nonlinear optical, array devices such as SEED arrays and bistable etalon arrays if and as when such devices are available and it is considered advantageous to employ them (such a "switch" from spatial light modulators would, however, almost certainly entail the upgrading to a more powerful laser).

(3) The extramural program, which has been under way for approximately two years, would be guided on a year-to-year basis under two fairly specific themes:

(a) This research should generally be directed towards the ultimate fabrication of novel optical device hardware and the hardware implementation of neural network circuitry, although a period of design studies and system characterizations will almost certainly be required.

(b) Wherever possible, the extramural research program should, if only partially, be of benefit to the research efforts expended internally and vice versa. In short, a certain degree of correlation should exist between the two programs. An optical device, for example, developed in either program could be made available to the 'partners' in these programs. Nevertheless, it remains important that DOCT should be free to choose the nature and specific goals of its own research while welcoming any advice or recommendations concerning research in either program given to DOCT by outside contractors.

(4) Both the intramural and extramural research programs should be directed as much as possible towards the development of specific neural network technology which is directly applicable to both present and future military and defense-related needs.

APPENDIX A : THE NEURAL NETWORK LITERATURE

The body of literature pertinent to the subject of neural network research is unusual in two ways:

(1) Until very recently, there existed no preferred journals for the publication of neural network research whether theoretical or experimental in nature. Furthermore, since the very character of neural network research is multi-disciplinary, articles have appeared in an extremely diverse mix of journals, ranging from psychology research publications through the biosciences to journals dedicated to the physical sciences or pure mathematics. Fortunately, some six journals have been created within the past two years which are devoted largely or exclusively to the publishing of neural network research results.

(2) Partly because of the multi-disciplinary nature of neural network research and also because the field went through a period of 'disfavor' when many researchers felt obliged to use "euphemisms" to have their work published, the present terminology of neural networks is characterized by an anomalously large number of synonymous or equivalent terms (e.g., neural networks go by the description artificial neural systems, artificial neural networks, neuro-computer, neural computing, parallel distributed processors, and connectionist machines). This proliferation of synonymous terminology can both confuse the non-expert and compound the difficulties involved in searching the literature for the appropriate material.

Because of these inherent difficulties, the neural network community has responded with comprehensive published bibliographies (many of which are diskette based) which transcend these two specific complications. Ironically, the availability of these resources make neural network literature searches both simple and straightforward. One of the best of these sources is Klimasauskas [1989] which, since 1987, has been republished with updates each year. This 600 page book contains approximately 4000 references, each cross referenced according to key words, author, and subject headings. A comprehensive glossary of neural network terminology is given in the DARPA [1988] publication.

The situation has also improved dramatically within the past few years with respect to the availability of textbooks and treatises on the subject of neural networks. Approximately fifteen such books were published in 1988/1989 alone (whereas, previously, only five or six existed). The texts by Wasserman [1989] and Dayhoff [1989] represent excellent introductions to the field with a minimum of complex mathematics (as do the articles by Hecht-Nielsen [1988] and by Lippmann [1987, 1989]). Books by Simpson [1990], Khanna [1990], Kohonen [1977, 1984], and Soucek [1989] offer a more advanced and mathematical treatment of the subject (although now 'dated', the books by Rosenblatt

[1962] and Minsky and Papert [1969] offer detailed treatments of much of the underlying mathematics of neural network theory). Some of the more philosophical and esoteric aspects of neural network technology and its potential are discussed in books by Johnson and Brown [1988], McCulloch [1965], and Moravec [1988]. The book edited by Anderson and Rosenfeld [1988] contains an extensive collection of original research papers back to the 1940's and offers an excellent introduction to the development of the field. The texts by Hoppensteadt [1986] and Koch and Segev [1989] present detailed, mathematical descriptions of realistic models of neuronal behavior. Compendia of research papers also offer excellent sources of material and include many conference proceedings (IEEE and SPIE conferences, nearly all of which have occurred since 1985), special journal issues (e.g., IEEE Computer, March 1988) as well as 'special topic' publications of selected research papers (e.g., that edited by Caulfield and Gheen [1989] on optical computing and optical neural networks).

REFERENCES

Aarts E. and Korst J., (1990), **Simulated Annealing and Boltzmann Machines**, John Wiley (New York).

Abu-Mostafa Y. S. and Psaltis D., (1987), "Optical neural computers", Sci. Amer. 256, no. 3 (March), pp. 88-95.

Ackley D. H., Hinton G. E., and Sejowski T. J., (1985), "A learning algorithm for Boltzmann machines", Cognitive Science 9, pp. 147-169.

Agranat A., Neugebauer C. F., and Yariv A., (1988), "Parallel optoelectronic realization of neural networks using CID technology", Appl. Optics 27, pp. 4354-4355.

Aleksander I. (Ed.), (1989), **Neural Computing Architectures : The Design of Brain-Like Machines**, MIT Press (Cambridge).

Amari S. I., (1967), "A theory of adaptive pattern classifiers", IEEE Trans. Electronic Computers EC-16, pp. 299-307.

Amari S. I., (1977), "Neural theory of association and concept-formation", Biological Cybernetics 26, pp. 175-185.

Anderson D. Z., (1987), "Adaptable interconnects for optical neuromorphs: Demonstration of a photorefractive projection operator", pp. 577-584 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Anderson D. Z. (Ed.), (1988), **Neural Information Processing Systems**, Amer. Institute of Physics (New York).

Anderson D. Z., (1989), "Optical systems that imitate human memory", Computers in Physics, Mar/Apr 1989, pp. 19-25.

Anderson J. A., (1972), "A simple neural network generating an interactive memory", Math. Biosciences 14, pp. 197-220.

Anderson J. A. and Rosenfeld E. (Eds.), (1988), **Neurocomputing - Foundations of Research**, MIT Press (Cambridge).

Arsenault H. H. and Macukow B., (1989), "Neural network model for fast learning and retrieval", Optical Eng. 28, pp. 506-512.

Athale R. A., (1989), "Optical implementations of neural computing", pp. 2-4 in **Optical Computing - 1989 Technical Digest Series** (vol. 9) (Optical Soc. Press, Washington, D.C., 1989).

Athale R. A. and Davis J. (Eds.), (1988), **Neural Network Models for Optical Computing**, Proc. SPIE 882 (SPIE Press, Bellingham, WA).

Athale R. A., Friedlander C. B., and Kushner C. B., (1986), "Attentive associative architectures and their implications to optical computing", Proc. SPIE 625, pp. 179-188.

Athale R. A., Szu H. H., and Friedlander C. B., (1986), "Optical implementation of associative memory with controlled nonlinearity in the correlation domain", Optics Lett. 11, pp. 482-484.

Barto A. G., Sutton R. S., and Anderson C. W., (1983), "Neuronlike adaptive elements that can solve difficult learning control problems", IEEE Trans. Systems, Man, and Cybernetics 13, pp. 834-846.

Baum E. B. and Haussler D., (1988), "What size net gives valid generalization?", Neural Computation 1, pp. 151-160.

Bayley J. S. and Fiddy M. A., (1987), "On the use of the Hopfield model for optical pattern recognition", Optics Comm. 64, pp. 105-110.

Bernstein J., (1981), "Profiles: Marvin Minsky", The New Yorker (December 14), pp. 50-126.

Block H. D., (1962), "The perceptron: A model for brain functioning. I.", Rev. Mod. Phys. 34, pp. 123-135.

Carpenter G. and Grossberg S., (1986), "Adaptive resonance theory: Stable self-organization of neurocognition codes in response to arbitrary lists of input patterns", pp. 45-62 in **Eighth Annual Conference of the Cognitive Science Society**, Lelbaum (Hillsdale, NJ).

Carpenter G. and Grossberg S., (1987), "ART2: Self-organization of stable category recognition codes for analog input patterns", Appl. Optics 26, pp. 4919-4930.

Carpenter G. and Grossberg S., (1988), "The ART of adaptive pattern recognition by a self-organizing neural network", IEEE Computer Magazine (March), pp. 77-88.

Casasent D. P. and Barnard E., (1990), "Adaptive-clustering optical neural net", Appl. Optics 29, pp. 2603-2615.

Cathey W. T., Wagner K., and Miceli W. J., (1989), "Digital computing with optics", Proc. IEEE 77, pp. 1558-1572.

Caudill M. and Butler C., (1987), **IEEE First International Conference on Neural Networks** (4 vol.),IEEE Press, San Diego.

Caulfield H. J., (1987), "Parallel N^4 weighted optical interconnections", pp. 595-597 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Caulfield H. J., (1989), "Computing with light", Byte 14, no. 10 (October), pp. 231-237.

Caulfield H. J. and Gheen G. (Eds.), (1989), **Selected Papers on Optical Computing**, vol. 1142 in SPIE Milestone Series (SPIE Optical Engineering Press, Bellingham, WA).

Caulfield H. J., Kinser J. and Rogers S. K., (1989), "Optical neural networks", Proc. IEEE 77, pp. 1573-1583.

Channin D. J., (1989), "Joining forces", Byte 14 (October), pp. 244-248.

Churchland P. M. and Churchland P. S., (1990), "Could a machine think?", Sci. Amer. 262 (January), pp. 32-37.

Cohen M. S., (1986), "Design of a new medium for volume holographic information processing", Appl. Optics 25, pp. 2288-2294.

Cohen M. A. and Grossberg S. G., (1983), "Absolute stability of global pattern formation and parallel memory storage by competitive neural networks", IEEE Trans. Systems, Man, and Cybernetics 13, pp.815-826.

Cotterill R. M. J. (Ed.), (1986), **Computer Simulation in Brain Science**, Cambridge Univ. Press (U.K.).

Cottrell D. M., Davis, J. A., Hedman T. R., and Lilly R. A., (1990), "Multiple imaging phase-encoded optical elements written programmable spatial light modulators", Appl. Optics 29, pp. 2505-2509.

Craft N. C. and Prise M. E., (1990), "Processor does light logic", Laser Focus (May), pp. 191-200.

DARPA, (1988), **DARPA Neural Network Study**, AFCEA International Press (Fairfax, Virginia).

Dayhoff J., (1989), **Neural Net Architectures - An Introduction**, Van Nostrand (New York).

Dunning G. J., Marom E., Owechko Y., and Soffer B. H., (1985), "All-optical associative holographic memory with feedback using phase conjugate mirrors", Proc. SPIE 625, pp. 205-213.

Farhat N. H., (1987), "Architectures for optoelectronic analogs of self-organizing neural networks", Optics Lett. 12, pp. 448-450.

Farhat N. H., Psaltis D., Prata A., and Paek E., (1985), "Optical implementation of the Hopfield model", Appl. Optics 24, pp. 1469-1475.

Farhat N. H. and Shae Z. Y., (1987), "Architectures and methodologies for self-organization and stochastic learning in opto-electronic analogs of neural nets", pp. 557-563 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Feldman J. A. and Ballard D. H., (1982), "Connectionist models and their properties", Cognitive Science 6, pp. 205-254.

Fox G. C. and Messina P. C., (1987), "Advanced computer architectures", Sci. Amer. 257 (October), pp. 67-74.

Frye R. C., Rietman E. A., Wong C. C., and Chin B. L., (1989), "An investigation of adaptive learning implemented in an optically controlled neural network", pp. 457-463 (vol. 2) in **IEEE International Joint Conference on Neural Networks** (Washington, D.C., 1989).

Geman S. and Geman D., (1984), "Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images", IEEE Trans. Pattern Analysis and Machine Intelligence 6, pp. 721-741.

Giles C. L. and Maxwell T., (1987), "Learning, invariance, and generalization in high-order neural networks", Appl. Optics 26, pp. 4972-4978.

Gindi G., Gmitro A., and Parthasarathy K., (1987), "Winner-take-all networks and associative memory: Analysis and optical realization", pp. 607-614 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Grossberg S., (1968a), "A prediction theory for some nonlinear functional-differential equations: II. Learning of patterns", J. Math. Analysis and Applications 22, pp. 490-522.

Grossberg S., (1968b), "Some nonlinear networks capable of learning a spatial pattern of arbitrary complexity", Proc. Nat. Acad. Sci. 59, pp. 368-372.

Grossberg S., (1969), "On learning, information, lateral inhibition, and transmitters", Math. Biosciences 4, pp. 255-310.

Grossberg S., (1976), "Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors", Biological Cybernetics 23, pp. 121-134.

Grossberg S., (1987), "Recent developments in the design of real-time non-linear neural network architectures", pp. 1-75 (vol. 1) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Grossberg S., (1988), "Nonlinear neural networks: principles, mechanisms, and architectures", Neural Networks 1, pp. 17-61.

Hebb D., (1949), **Organization of Behavior**, John Wiley and Sons (New York).

Hecht-Nielsen R., (1987), "Neural network nearest matched-filter classification of spatiotemporal patterns", Appl. Optics 26, pp. 1892-1899.

Hecht-Nielsen R., (1988), "Neurocomputing: picking the human brain", IEEE Spectrum, March 1988, pp. 36-41.

Hinton G. E., (1989), "Connectionist learning procedures", Artificial Intelligence 40, pp. 185-234.

Hinton G. E. and Sejnowski T. J., (1986), "Learning and relearning in Boltzmann machines", pp. 282-317 in **Parallel Distributed Processing** (Vol. 1), MIT Press (Cambridge).

Hopfield J. J., (1982), "Neural networks and physical systems with emergent collective computational abilities", Proc. Nat. Acad. Sci. 79, pp. 2554-2558.

Hopfield J. J., (1984), "Neurons with graded response have collective computational properties like those of two-state neurons", Proc. Nat. Acad. Sci. 81, pp. 3088-3092.

Hopfield J., (1987), "Networks, computations, logic and noise", pp. 107-141 (vol. 1) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Hoppensteadt F. C., (1986), **An Introduction to the Mathematics of Neurons**, Cambridge University Press (U.K.).

Jang J.-S., Jung S.-W., Lee S.-Y., and Shin S.-Y., (1988), "Optical implementation of the Hopfield model for two-dimensional associative memory", Optics Lett. 13, pp. 248-250.

Jang J.-S., Shin S.-Y., and Lee. S.-Y., (1988), "Optical implementation of quadratic associative memory with outer-product storage", Optics Lett. 13, pp. 693-695.

Johnson K. M., Kranzdorf M., Bigner B. J., and Zhang L., (1989), "Optical associative memory utilizing electrically and optically addressed liquid crystal spatial light modulators", pp. 32-35 in **Optical Computing - 1989 Technical Digest Series** (vol. 9) (Optical Soc. Press, Washington, D.C., 1989).

Johnson R. C. and Brown C., (1988), **Cognizers : Neural Networks and Machines That Think**, John Wiley (New York).

Kalil R. E., (1989), "Synapse formation in the developing brain", Sci. Amer. 261, no. 6 (December), pp. 76-85.

Khanna T., (1990), **Foundations of Neural Networks**, Addison-Wesley Pub. Co. (Reading, Massachusetts).

Kinser J. M., Caulfield H. J., and Shamir J., (1988), "Design for a massive all-optical bidirectional associative memory: the big BAM", Appl. Optics 27, pp. 3442-3444.

Kirkpatrick S., Gelatt Jr. C. D., and Vecchi M. P., (1983), "Optimization by simulated annealing", Science 220, pp. 671-680.

Kitayama K., Yoshinaga H., and Hara T., (1989), "Experiments of learning in optical perceptron-like and multilayer neural networks", pp. 465-471 (vol. 2) in **IEEE International Joint Conference on Neural Networks** (Washington, D.C., 1989).

Klimasauskas C. C. (Ed.), (1989), **The 1989 Neuro-Computing Bibliography**, MIT Press (Cambridge).

Koch C. and Segev I. (Eds.), (1989), **Methods in Neuronal Modeling - From Synapses to Networks**, MIT Press (Cambridge).

Kohonen T., (1972), "Correlation matrix memories", IEEE Trans. Computers 21, pp. 353-359.

Kohonen T., (1977), **Associative Memory : A System-Theoretical Approach**, Springer (Berlin).

Kohonen T., (1982), "Self-organized formation of topologically correct feature maps", Biological Cybernetics 43, pp. 59-69.

Kohonen T., (1984), **Self-organization and Associative Memory**, Vol. 8 in Series in Information Sciences, Springer (Berlin).

Kohonen T., (1987), "State of the art in neural computing", pp. 77-90 (vol. 1) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Kohonen T., (1988), "An introduction to neural computing", Neural Networks 1, pp. 3-16.

Kosko B., (1987), "Adaptive bidirectional associative memories", Appl. Optics 26, pp. 4947-4960.

Lalanne Ph., Taboury J., and Chavel P., (1987), "A proposed generalization of Hopfield's algorithm", Optics Comm. 63, pp. 21-25.

Lin F., Savant G., Brown R., Duo C.-P., and Jannson T., (1989), "Highly paralle! optical adaptive neural computers with layered and segmented structures", pp. 487-494 (vol. 2) in **IEEE International Joint Conference on Neural Networks** (Washington, D.C., 1989).

Lippmann R., (1987), "An introduction to computing with neural nets", IEEE ASSP Magazine 4 (no. 2), pp. 4-22.

Lippmann R. P., (1989), "Review of neural networks for speed recognition", Neural Computation 1, pp. 1-38.

McCulloch W., (1965), **Embodiments of Mind**, MIT Press (Cambridge).

McCulloch W. S. and Pitts W., (1943), "A logical calculus of the ideas immanent in nervous activity", Bulletin Math. Biophysics 5, pp. 115-133.

Mead C., (1987), "Silicon models of neural computation", pp. 91-106 (vol. 1) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Mead C., (1987), **Analog VLSI and Neural Systems**, Addison-Wesley (New York).

Miller R. K., (1987), **Neural Networks - Implementing Associative Memory Models in Neurocomputers** (2 vol.), SEAI Technical Publications (Madison, GA).

Minsky M., (1961), "Steps towards artificial intelligence", Proc. IRE 49, pp. 5-30.

Minsky M. and Papert S., (1969), **Perceptrons**, MIT Press (Cambridge).

Moravec H., (1988), **Mind Children - The Future of Robot and Human Intelligence**, Harvard Univ. Press, Mass., 1988.

Ohta J., Oita M., Tai S., Hara K., and Kyuma K., (1990), "Opto-electronic implementation of a large-scale neural network using multiplexing techniques", Trans. IEICE E73, pp. 41-45.

Owechko Y., (1987), "Optoelectronic resonator neural networks", Appl. Optics 26, pp. 5104-5111.

Owechko Y., Dunning G. J., Marom E., and Soffer B. H., (1987), "Holographic associative memory with nonlinearities in the correlation domain", Appl. Optics 26, pp. 1900-1910.

Parker D., (1982), "Learning logic", Invention report S81-64, File 1, Office of Technology Licensing, Stanford University.

Parker D., (1987), "Optimal algorithms for adaptive networks: Second order back propagation, second order direct propagation, and second order Hebbian learning", Proc. IEEE First Int. Conf. on Neural Networks, Vol. II, pp. 593-600 (IEEE Press, San Diego).

Peled A., (1987), "The next computer revolution", Sci. Amer. 257 (October), pp. 57-64.

Peterson C., Redfield S., Keeler J. D., and Hartman E., (1990), "Optoelectronic implementation of multilayer neural networks in a single photorefractive crystal", Optical Eng. 29, pp. 359-367.

Pineda R. J., (1988), "Generalization of backpropagation to recurrent and higher order networks", pp. 602-611 in **Neural Information Processing Systems**, D. Z. Anderson (ed.), American Institute of Physics (New York).

Pitts W. and McCulloch W. S., (1947), "How we know universals: The perception of auditory and visual forms", Bulletin Math. Biophysics 9, pp. 127-147.

Prise M. E., Streibl N., and Downs M. M., (1988), "Optical considerations in the design of digital optical computers", Optical and Quantum Electronics 20, pp. 49-77.

Psaltis D., Brady D., and Wagner K., (1988), "Adaptive optical networks using photorefractive crystals", Appl. Optics 27, pp. 1752-1759.

Psaltis D. and Farhat N., (1984), "A new approach to optical information processing based on the Hopfield model", pp. 24-25 in **Optics in Modern Science and Technology** (Sapporo, Japan, 1984).

Psaltis D. and Farhat N., (1985), "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback", Optics Lett. 10, pp. 98-100.

Psaltis D., Park C. H., and Hong J., (1988), "Higher order associative memories and their optical implementations", Neural Networks 1, pp. 149-163.

Psaltis D., Wagner K., and Brady D., (1987), "Learning in optical neural computers", pp. 549-555 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Psaltis D., Yamamura A. A., Hsu K., Lin S., Gu X., and Brady D., (1989), "Optoelectronic implementations of neural networks", IEEE Communications Magazine (November), pp. 37-48.

Rice R. A., Li W., and Moddel G., (1989), "High-speed optically-addressed spatial light modulator for optical computing", pp. 64-66 in **Optical Computing - 1989 Technical Digest Series** (vol. 9) (Optical Soc. Press, Washington, D.C., 1989).

Rosenblatt F., (1958), "The perceptron: A probabilistic model for information storage and organization in the brain", Psychological Review 65, pp. 386-408.

Rosenblatt F., (1962), **Principles of Neurodynamics**, Spartan Books (Washington, D.C.).

Rosenfeld E., (1987), "Neurocomputing - A new industry", pp. 831-838 (vol. 4) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Rumelhart D. E., Hinton G. E., and Williams R. J., (1986b), "Learning internal representations by back-propagating errors", Nature 323, pp. 533-536.

Rumelhart D. E., Hinton G. E., and Williams R. J., (1986a), "Learning internal representations by error propagation", pp. 318-362 in **Parallel Distributed Processing: Explorations in the Microstructures of Cognition** (Vol. 1), MIT Press (Cambridge).

Russell S., (1913), "A practical device to simulate the working of nervous discharges", J. Animal Behavior 3, pp. 15 - ff.

Sage J. P., (1989), "Electronic vs. optical implementations of neural networks", pp. 5-8 in **Optical Computing - 1989 Technical Digest Series** (vol. 9) (Optical Soc. Press, Washington, D.C., 1989).

Searle J. R., (1990), "Is the brain's mind a computer program?", Sci. Amer. 262 (January), pp. 26-31.

Sejnowski T. J. and Rosenberg C. R., (1987), "Parallel networks that learn to pronounce English text", Complex Systems 1, pp. 145-168.

Shamir J., Caulfield H. J., and Johnson R. B., (1989), "Massive holographic interconnection networks and their limitations", Appl. Optics 28, pp. 311-324.

Shynk J. J., (1990), "Performance surfaces of a single-layer perceptron", IEEE Trans. Neural Networks 1, pp. 268-274.

Simpson P. K., (1990), **Artificial Neural Systems - Foundations, Paradigms, Applications, and Implementations**, Pergamon Press (New York).

Soucek B., (1989), **Neural and Concurrent Real-Time Systems : The Sixth Generation**, John Wiley (New York).

Steinbuch K., (1961), "Die lernmatrix", Kybernetik 1, pp. 36-45.

Steinbuch K. and Widrow B., (1965), "A critical comparison of two kinds of adaptive classification networks", IEEE Trans. Electronic Computers 14, pp. 737-740.

Stornetta W. and Huberman B., (1987), "An improved three-layer backpropagation algorithm", Proc. IEEE First Int. Conf. on Neural Networks, Vol. II, pp. 637-634 (IEEE Press, San Diego).

Tank D. W. and Hopfield J. J., (1987), "Collective computation in neuronlike circuits", Sci. Amer. 257 (December), pp. 104-114.

Te Kolste R. D. and Guest C. C., (1987), "Optical competitive neural network with optical feedback", pp. 625-629 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Ticknor A. J. and Barrett H. H., (1987), "Optical implementations in Boltzmann machines", Optical Eng. 26, pp. 16-21.

Treleaven P., Pacheco M., and Vellasco M., (1989), "VLSI architectures for neural networks", IEEE Micro 9, no. 6 (December), pp. 8-27.

Uttley A., (1966), "The transmission of information and the effect of local feedback in theoretical and neural networks", Brain Research 2, pp. 21-50.

Verleysen M. and Jespers P. G. A., (1989), "An analog VLSI implementation of Hopfield's neural network", IEEE Micro 9 (December), pp. 46-55.

VonLehmen A., Paek E. G., Carrion L. C., Patel J. S., and Marrakchi A., (1989), "Optical implementation of a high order associative memory", pp. 473-475 (vol. 2) in **IEEE International Joint Conference on Neural Networks** (Washington, D.C., 1989).

Wagner K., Giles L., and Psaltis D., (1990), "Workshop considers optical neural networks", Laser Focus (August), pp. 127-130.

Wagner K. and Psaltis D., (1987), "Nonlinear etalons in competitive optical learning networks", pp. 585-594 (vol. 3) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Wasserman P. D., (1989), **Neural Computing - Theory and Practice**, Van Nostrand Reinhold (New York).

Werbos P., (1974), "Beyond regression: New tools for prediction and analysis in the behavioral sciences", Ph. D. Dissertation, Harvard University.

Widrow B., (1959), "Adaptive sample-data systems - A statistical theory of adaptation", WESCON 1959 Convention Record, Part 4, pp. 74-85.

Widrow B., (1987), "Adaline and Madaline - 1963", pp. 143-157 (vol. 1) in **IEEE First International Conference on Neural Networks** (IEEE, San Diego, 1987).

Widrow B. and Hoff M., (1960), "Adaptive switching circuits", WESCON 1960 Convention Record, Part 4, pp. 96-104.

Wilde J. and Hesselink L., (1989), "Implementation of dynamic Hopfield-like networks using photorefractive crystals", pp. 10-13 in **Optical Computing - 1989 Technical Digest Series** (vol. 9) (Optical Soc. Press, Washington, D.C., 1989).

Xu H., Yuan Y., Yu Y., Xu K., and Xu Y., (1990), "Performance of real time associative memory using a photorefractive crystal and liquid crystal eletrooptic switches", Appl. Optics 29, pp. 3375-3379.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

COMMUNICATIONS RESEARCH CENTRE
3701 CARLING AVE.
SHIRLEY'S BAY K2H 8S2

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

NEURAL AND OPTICAL NEURAL NETWORKS (U)

4. AUTHORS (Last name, first name, middle initial)

SALA, KENNETH L.

5. DATE OF PUBLICATION (month and year of publication of document)

FEBRUARY 1991

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

76

6b. NO. OF REFS (total cited in document)

145

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

TECHNICAL REPORT

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
SHIRLEY'S BAY K1A 0Z4

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

N/A

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

CRC REPORT 1434

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

N/A

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

UNLIMITED

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

THIS REPORT PRESENTS AN OVERVIEW OF NEURAL AND OPTICAL NEURAL NETWORK TECHNOLOGIES WHICH INCLUDES DETAILED DESCRIPTIONS OF SOME OF THE MORE IMPORTANT PARADIGMS SUCH AS SINGLE-AND MULTI-LAYER PERCEPTIONS, BACKPROPAGATION, HOPFIELD NETWORKS, ASSOCIATIVE MEMORIES, AND STOCHASTIC LEARNING. THESE PARADIGMS HAVE BEEN SHOWN TO HAVE PARTICULAR SUITABILITY FOR OPTICAL IMPLEMENTATIONS. OPTICAL NEURAL NETWORKS ARE DISCUSSED FROM THE POINT OF VIEW FO TWO CLASSES OF IMPLEMENTATIONS, NAMELY, (A) THOSE BASED UPON THE USE OF LINEAR OR NONLINEAR, TWO-DIMENSIONAL OPTICAL ELEMENTS SUCH AS SPATIAL LIGHT MODULATORS OR BISTABLE ETALON ARRAYS AND, (B) THOSE BASED UPON THE USE OF HOLOGRAPHIC ELEMENTS SUCH AS PHOTOREFRACTIVE CRYSTALS WHICH PERMIT A FULL THREE-DIMENTIONAL IMPLEMENTATION OF NUERAL NETWORK ARCHITECTURES. THIS REPORT CONCLUDES BY OFFERING A SERIES OF GENERAL CONCLUSIONS REGARDING THE PRESENT AND FUTURE DIRECTIONS FOR NEURAL AND OPTICAL NEURAL NETWORK RESEARCH AS WELL AS PRESENTING SPECIFIC RECOMMENDATIONS FOR THE CONDUCT OF AN INTRAMURAL/EXTRAMURAL RESEARCH PROGRAM INTO OPTICAL NEURAL NETWORKS UNDER THE GUIDANCE OF THE DIRECTORATE OF OPTICAL COMMUNICATIONS AND ELECTROPHOTONICS TECHNOLOGY OF THE COMMUNICATIONS RESEARCH CENTRE.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

NEURAL NETWORKS

OPTICAL NEURAL NETWORKS